



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM INOVAÇÃO TERAPÊUTICA

RHAYANNE THAÍS DE MORAES RAMOS

**DESENVOLVIMENTO DE SISTEMA DE ANÁLISE DE IMAGENS PARA
QUANTIFICAÇÃO DE METABÓLITOS SECUNDÁRIOS A PARTIR DE
FINGERPRINTS POR CROMATOGRAFIA EM CAMADA DELGADA**

Recife

2019

RHAYANNE THAÍS DE MORAES RAMOS

**DESENVOLVIMENTO DE SISTEMA DE ANÁLISE DE IMAGENS PARA
QUANTIFICAÇÃO DE METABÓLITOS SECUNDÁRIOS A PARTIR DE
FINGERPRINTS POR CROMATOGRAFIA EM CAMADA DELGADA**

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Inovação Terapêutica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção de título de mestre em inovação terapêutica.

Área de concentração: Desenho, Modelagem Molecular e Preparação de Produtos Bioativos.

Orientador: Prof. Dr. Luiz Alberto Lira Soares

Coorientador (es): Prof. Dr. Wellington Pinheiro dos Santos
Profa. Dra. Magda Rhayanny A. Ferreira

Recife

2019

Catálogo na fonte:
Bibliotecária Claudina Queiroz, CRB4/1752

Ramos, Rhayanne Thaís de Moraes

Desenvolvimento de sistema de análise de imagens para quantificação de metabólitos secundários a partir de *fingerprints* por cromatografia em camada delgada / Rhayanne Thaís de Moraes Ramos - 2019.

120 folhas: il., fig., tab.

Orientador: Luiz Alberto Lira Soares

Coorientadores: Wellington Pinheiro dos Santos

Magda Rhayanny A. Ferreira

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-Graduação em Inovação Terapêutica. Recife, 2019.

Inclui referências e anexo.

1. Cromatografia 2. Plantas Medicinais 3. Polifenóis

I. Soares, Luiz Alberto Lira (Orient.) II. Santos, Wellington Pinheiro dos (Coorient.) III. Ferreira, Magda Rhayanny A.(Coorient.) IV. Título

615.3

CDD (22.ed.)

UFPE/CB-2019-365

RHAYANNE THAÍS DE MORAES RAMOS

**DESENVOLVIMENTO DE SISTEMA DE ANÁLISE DE IMAGENS PARA
QUANTIFICAÇÃO DE METABÓLITOS SECUNDÁRIOS A PARTIR DE
FINGERPRINTS POR CROMATOGRAFIA EM CAMADA DELGADA**

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Inovação Terapêutica da Universidade Federal de Pernambuco, como requisito parcial para a obtenção de título de mestre em inovação terapêutica.

Aprovado em: 29/ 08/ 2019.

BANCA EXAMINADORA

Prof. Dr. Luiz Alberto Lira Soares (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Marcelo Zaldini Hernandes (Membro Interno)
Universidade Federal de Pernambuco

Prof. Dr. Ricardo Oliveira da Silva (Membro Externo)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Agradeço primeiramente a Deus pelas oportunidades que me foram dadas e por ter me dado forças e iluminado meu caminho para que pudesse seguir mais uma etapa da minha vida. Agradeço aos meus pais, em especial minha mãe Tereza por ser a base da minha vida, pelo amor, orações, incentivo e apoio para que eu pudesse alcançar todos os meus objetivos. A minha irmã Yasmim pelos conselhos e exemplo. Agradeço a Allan pelo companheirismo e por ter tornado minha jornada mais prazerosa.

Ao meu orientador Luiz Soares pela oportunidade, confiança, e todos os ensinamentos passados durante todos estes anos de trabalho. Considero como grande exemplo profissional a ser seguido.

Ao professor Wellington pela paciência de ter-me ajudado e por me acrescentar novos conhecimentos, sua contribuição foi essencial para o desenvolvimento desse trabalho. À Magda por me incentivar na busca do conhecimento e por me estimular nessa jornada, pelas boas conversas, dicas e sugestões, pelos ensinamentos e pela dedicação dispensados durante toda minha fase de aprendizado. À Sarah por todos conselhos e paciência, por ter dedicado uma parte do seu tempo para me ajudar com conselhos e sugestões, além da amizade que fez minha rotina mais prazerosa. Meus agradecimentos vão também para Patrícia e Maíra que me ajudaram no desenvolvimento desse trabalho, sem a ajuda de vocês isso não teria sido possível. Agradeço imensamente a todos colaboradores do Laboratório de Farmacognosia por acrescentar na minha vida com ensinamentos e pela amizade.

As minhas amigas Amanda, Juliane, Marília e Mônica por todo apoio e incentivo desde o tempo da graduação. Agradeço pelos conselhos, paciência e amizade desenvolvida ao longo desses anos.

Agradeço a todas as pessoas que de alguma forma me ajudaram nesse processo, citadas aqui ou não, algumas mais de perto, outras nem tanto, porém todas com a sua devida importância e parcela de contribuição.

RESUMO

A relevância do estudo de plantas medicinais reside nas inúmeras aplicabilidades no campo farmacêutico, cuja diversidade química assegura o emprego destas matrizes biológicas tanto como fontes de diversos compostos relevantes quanto como papel de insumo farmacêutico de ativo natural (IFAN). Entre os diferentes grupos de compostos do metabolismo secundário vegetal, aos polifenóis é atribuída destacada relevância pelo desempenho de diversas propriedades biológicas. Porém, em razão da diversidade estrutural deste grupo de compostos, as análises qualitativas e quantitativas dos polifenóis exigem o emprego de técnicas de elevadas complexidades e custo operacional significativo. Assim, o desenvolvimento de novas abordagens que possibilitem a análise (qualitativa/quantitativa) de compostos naturais tem papel importante para o avanço dos métodos de análise e controle de qualidade de toda a cadeia de plantas medicinais e fitoterápicos. Neste contexto, o objetivo desse trabalho é configurar redes neurais artificiais (RNAs) como estratégia analítica inovadora, simples e de baixo custo para a quantificação de polifenóis em matrizes vegetais. Para o desenvolvimento e avaliação da abordagem analítica, foram empregados diversos padrões representativos dos polifenóis (ácido gálico, quercetina, rutina, ácido cafeico e catequina) e análises foram realizadas em *Schinus terebinthifolius* e *Anadathera colubrina*, cujas espécies apresentam potencial farmacológico. Para tanto, curvas de calibração dos padrões foram preparadas por cromatografia em camada delgada (CCD) para posterior registro e análise de imagens; e por cromatografia líquida de alta eficiência (CLAE-DAD), cuja finalidade foi de verificar a relação entre os dados obtidos. As cromatoplasmas foram digitalizadas sob os comprimentos de onda de 254 nm e 366nm e as imagens foram tratadas pela ferramenta de imagem ImageJ e Sidtermo, gerando atributos de imagem para alimentar o sistema de rede. Diversos modelos de Kernel foram testados com a finalidade de encontrar a configuração que apresentasse maior desempenho para modelar os dados e assim associar os atributos de imagem com a respectiva concentração. Os modelos de Kernel que apresentaram maior desempenho quando o sistema foi alimentado pelos atributos de imagem gerados pelo ImageJ para a maioria dos compostos testados foram o linear, dilatação e erosão. Os atributos de imagem gerados pela ferramenta de imagem Sidtermo ao serem correlacionados com a concentração dos marcadores de interesse evidenciaram maior precisão de

resultados, devido ao menor erro percentual médio encontrado para a maioria das análises realizadas.

Palavras chave: Análise de imagem. Cromatografia. Plantas Medicinais. Polifenóis. Qualidade. Rede Neural Artificial.

ABSTRACT

The importance of the study of medicinal plants lies in the several applications in the pharmaceutical field, whose chemical diversity ensure the use of such biological matrices both as source of relevant compounds and as natural active pharmaceutical ingredient (nAPI). Among the groups of compounds from the herbal metabolism, the polyphenols are plays a pivotal role on the biological performance of several species. However, due to the structural variability of the group of compounds, the qualitative and quantitative of polyphenols, requires the use of high complex techniques and significant operational cost. Thus, the development of new approaches that allow the analysis (qualitative / quantitative) of natural compounds plays an important role in advancing methods of analysis and quality control of a whole series of medicinal plants and phytopharmaceuticals. In this way, the aim of this study is to set up artificial neural networks (ANN) as innovative, simple and low cost strategy in order to quantify polifenolic compounds in herbal matrices.

Thus, several representative standards of polyphenols (gallic acid, quercetin, rutin, caffeic acid and catechin) were used to develop and evaluate the analytical procedure, and the validated protocol was performed to evaluate samples of *Schinus terebinthifolius* and *Anadathera colubrina*. Regarding the calibration curves of the standards, the thin layer chromatogram were prepared, followed by recording and analysis of images; and, verified its relationship with the quantitative data by high performance liquid chromatography (HPLC).. The TLC-plates were digitized under the wavelengths of 254 nm and 366 nm and the images were treated by the tools *image J* and *Sidtermo*, generating image attributes to feed the network system. Several kernel models were tested for the purpose of finding a configuration that presented higher performance to model the data and thus associate image attributes with attention. The Kernel models that are most important for most of the tests tested were linear, dilatation, and erosion. The image attributes generated by the *Sidtermo* image tool when correlated with the concentration of the markers of interest evidenced greater accuracy of results, due to the lower average percentage error found for most analyses Carried out.

.

Keywords: Image analysis. Chromatography. Medicinal Plants. Polyphenols. Quality. Artificial Neural Network.

LISTA DE FIGURAS

Figura 1 –	Representação da estrutura dos taninos hidrolisáveis	28
Figura 2 –	Representação da estrutura dos taninos condensados.....	29
Figura 3 –	Representação da estrutura química do flavonoide de quercetina.	30
Figura 4 –	Representação das estruturas químicas dos principais ácidos fenólicos.....	30
Figura 5 –	Representação de rede neural artificial.....	36
Figura 6–	Quatro sentidos da vizinhança definido para o cálculo de os recursos de textura Haralick.....	40
Figura 7 –	Cromatoplasmas desenvolvidas para análise de imagem do padrão ácido gálico. Faixa de concentração 20-30 µg/mL. Sistema de eluição - tolueno: AcOEt: ácido fórmico:MeOH (3:3:0,8:0,2). A – Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B - Cromatoplasma visualizada sob o comprimento de onda de 366 nm.....	52
Figura 8 –	Cromatoplasmas desenvolvidas para análise de imagem do padrão quercetina. Faixa de concentração: 7,5-17,5 µg/mL. Sistema de eluição - tolueno: AcOEt: HCOOH (6:4:0,3). A – Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B – Cromatoplasma visualizada sob o comprimento de onda de 366 nm	53
Figura 9 –	Cromatoplasmas desenvolvidas para análise de imagem do padrão Rutina. Faixa de concentração: 15-27 µg/mL. Sistema de eluição - AcOEt: HCOOH: H2O (90:5:5). A – Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B - Cromatoplasma visualizada sob o comprimento de onda de 366 nm	54
Figura 10 –	Cromatoplasmas desenvolvidas para análise de imagem do padrão ácido cafeico. Faixa de concentração: 25-35 µg/mL. Sistema de eluição - Tolueno: AcOEt: HCOOH (4,5:3:0,2). A -Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B – Cromatoplasma visualizada sob o comprimento de onda de 366 nm.....	55

Figura 11 –	Cromatoplasmas desenvolvidas para análise de imagem do padrão Catequina. Faixa de concentração: 12-20 µg/mL. Sistema de eluição - tolueno: AcOEt: HCOOH: MeOH (3:3:0,8:0,2). A - Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B - Cromatoplasma visualizada sob o comprimento de onda de 366 nm	56
Figura 12 –	Cromatoplasmas desenvolvidas para análise de imagem da fração acetato de etila de <i>S. terebinthifolius</i> . Faixa de concentração: 30-40µg/mL. Sistema de eluição - tolueno: AcOEt: HCOOH: MeOH (3:3:0,8:0,2). A – Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B – Cromatoplasma visualizada sob o comprimento de onda de 366 nm. Padrões à direita da placa: ácido gálico; ácido cafeico e quercetina	57
Figura 13 –	Cromatoplasmas desenvolvidas para análise de imagem da fração acetato de etila de <i>A. colubrina</i> . Faixa de concentração: 300-400 µg/mL. Sistema de eluição - tolueno: AcOEt: HCOOH: MeOH (3:3:0,8:0,2). A – Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B – Cromatoplasma visualizada sob o comprimento de onda de 366 nm. Padrão à direita da placa catequina.....	59
Figura 14 –	Imagem de cromatoplasma do padrão ácido gálico após processamento no software ImageJ.....	59
Figura 15 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 254 nm.....	64
Figura 16 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 254 nm.....	64
Figura 17 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 366 nm.....	65

Figura 18 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 366 nm.	66
Figura 19 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 254 nm.....	68
Figura 20 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 254 nm.....	68
Figura 21 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 366 nm.....	69
Figura 22 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 366 nm.....	70
Figura 23 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão quercetina em 254 nm.	71
Figura 24 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão quercetina em 254 nm.....	72
Figura 25 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão quercetina em 366 nm.	73
Figura 26 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão quercetina em 366 nm.	73
Figura 27 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão quercetina em 254 nm.....	75
Figura 28 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão quercetina em 254 nm.....	75

Figura 29 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão quercetina em 366 nm.	77
Figura 30 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão quercetina em 366 nm.	77
Figura 31 –	Boxplot das imagens tratadas pelo ImageJ para o teste de divisão percentual dos modelos analisados para o padrão rotina em 254 nm.....	79
Figura 32 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão rotina em 254 nm.....	79
Figura 33 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão rotina em 366 nm.	80
Figura 34 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão rotina em 366 nm.	81
Figura 35 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão rotina em 254 nm.....	82
Figura 36 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão rotina em 254 nm.....	83
Figura 37 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão rotina em 366 nm.	84
Figura 38 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão rotina em 366 nm.....	84
Figura 39 –	Boxplot das imagens tratadas pelo ImageJ para o teste de divisão percentual dos modelos analisados para o padrão rotina em 254 nm.	86

Figura 40 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 254 nm.	86
Figura 41 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão ácido cafeico em 366 nm.	87
Figura 42 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 366 nm.	88
Figura 43 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido cafeico em 254 nm.	89
Figura 44 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 254 nm.	91
Figura 45 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido cafeico em 366 nm.	91
Figura 46 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 366 nm.	93
Figura 47 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão catequina em 254 nm.	93
Figura 48 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão catequina em 254 nm.	94
Figura 49 –	Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão catequina em 366 nm.	95
Figura 50 –	Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão catequina em 366 nm.	95

Figura 51 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão catequina em 254 nm.	96
Figura 52 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão catequina em 254 nm.	97
Figura 53 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão catequina em 366 nm.	98
Figura 54 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão catequina em 366 nm.	98
Figura 55 –	Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão catequina em 366 nm.	100
Figura 56 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de <i>S. terebinthifolius</i> em 366 nm.....	101
Figura 57 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de <i>A. colubrina</i> em 254 nm.	103
Figura 58 –	Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de <i>A. colubrina</i> em 366 nm.	104
Figura 59 –	Gráficos de resíduos obtidos para as curvas de calibração dos padrões de polifenóis estudados.....	107
Figura 60 –	Gráficos de resíduos obtidos para as curvas de calibração de ácido gálico em <i>S. terebinthifolia</i> (A) e catequina em <i>A. colubrina</i> (B)	108
Figura 61 –	Curvas de calibração obtidas por CLAE dos padrões de polifenóis estudados.	108
Figura 62 –	Curvas de calibração obtidas por CLAE de ácido gálico em <i>S. terebinthifolia</i> (A) e catequina em <i>A. colubrina</i> (B).....	109

Figura 63 – Cromatograma obtido por CLAE da fração acetato de etila de <i>S. terebinthifolius</i> em 280 nm.....	110
Figura 64 – Cromatograma obtido por CLAE do extrato bruto de <i>A. colubrina</i> em 280 nm.	110

LISTA DE QUADROS

Quadro 1 –	Concentrações e sistemas de eluição para os respectivos padrões empregados para preparação de cromatoplasmas.....	41
Quadro 2 –	Condições de imagem utilizadas no ImageJ para análise.....	46
Quadro 3 –	Medidas de imagem selecionadas para análise.	46
Quadro 4 –	Condições cromatográficas utilizadas nas análises por CLAE..	49

LISTA DE TABELAS

Tabela 1 –	Valores do coeficiente de determinação (R ²) para o padrão ácido gálico ao aplicar relação linear. Faixa de concentração: 20-30 µg/mL.....	60
Tabela 2 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 254 nm com atributos de imagem gerados pelo ImageJ	62
Tabela 3 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 366 nm com atributos de imagem gerados pelo Image J	65
Tabela 4 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 254 nm com atributos de imagem gerados pelo Sidtermo	67
Tabela 5 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 366 nm com atributos de imagem gerados pelo Sidtermo.....	69
Tabela 6 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 254 nm com atributos de imagem gerados pelo ImageJ	70
Tabela 7 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 366 nm com atributos de imagem gerados pelo ImageJ.....	72
Tabela 8 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 254 nm dos atributos de imagem gerados pelo Sidtermo.....	74
Tabela 9 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 366 nm dos atributos de imagem gerados pelo Sidtermo.....	76
Tabela 10 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rutina em 254 nm com atributos de imagem gerados pelo ImageJ.....	78

Tabela 11 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 366 nm com atributos de imagem gerados pelo ImageJ.....	80
Tabela 12 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 254 nm dos atributos de imagem gerados pelo Sidtermo.....	82
Tabela 13 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 366 nm dos atributos de imagem gerados pelo Sidtermo.....	83
Tabela 14 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 254 nm com atributos de imagem gerados pelo ImageJ.....	85
Tabela 15 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 366 nm com atributos de imagem gerados pelo ImageJ.....	87
Tabela 16 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 254 nm dos atributos de imagem gerados pelo Sidtermo.....	88
Tabela 17 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 366 nm dos atributos de imagem gerados pelo Sidtermo.....	90
Tabela 18 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 254 nm com atributos de imagem gerados pelo ImageJ.....	92
Tabela 19 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 366 nm com atributos de imagem gerados pelo ImageJ.....	94
Tabela 20 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 254 nm dos atributos de imagem gerados pelo Sidtermo.....	95
Tabela 21 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 366 nm dos atributos de imagem gerados pelo Sidtermo.....	97

Tabela 22 –	Valores de erro percentual médio (EPM) para cada modelo analisado para <i>S. terebinthifolius</i> em 254 nm dos atributos de imagem gerados pelo Sidtermo.....	99
Tabela 23 –	Valores de erro percentual médio (EPM) para cada modelo analisado para <i>S. terebinthifolius</i> em 366 nm dos atributos de imagem gerados pelo Sidtermo.....	101
Tabela 24 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 254 nm dos atributos de imagem gerados pelo Sidtermo.....	103
Tabela 25 –	Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 366 nm dos atributos de imagem gerados pelo Sidtermo.....	104
Tabela 26 –	Avaliação estatística da resposta de linearidade de soluções dos padrões polifenólicos e amostras de <i>S. terebinthifolius</i> e <i>A. colubrina</i>	106

LISTA DE ABREVIATURAS E SIGLAS

ANVISA -	Agência Nacional de Vigilância Sanitária
CCD -	Cromatografia em Camada Delgada
CCD-AE -	Cromatografia em Camada Delgada de Alta Eficiência
CLAE-DAD -	Cromatografia Líquida de Alta Eficiência acoplada a Detector de Arranjo de Fotodiodos
CV -	Coefficiente de Variação
DPR -	Desvio Padrão Relativo
EPM -	Erro percentual médio
EQM -	Erro Quadrático Médio
HPLC -	<i>High Performance Liquid Chromatography</i>
RNA -	Rede Neural Artificial

SUMÁRIO

1	INTRODUÇÃO.....	23
1.1	OBJETIVOS	25
1.1.1	OBJETIVO GERAL.....	25
1.1.2	OBJETIVOS ESPECÍFICOS.....	25
2	FUNDAMENTAÇÃO TEÓRICA.....	26
2.1	POLIFENOIS.....	26
2.2	<i>Schinus terebinthifolius</i>	30
2.3	<i>Anadenanthera colubrina</i>	32
2.4	TÉCNICAS ANALÍTICAS DE CONTROLE DE QUALIDADE	33
2.5	QUIMIOMETRIA.....	34
2.5.1	Rede Neural Artificial (RNA)	35
2.5.1.1	Máquina de Aprendizado Extremo (Extreme Learning Machine - ELM).....	36
2.6	ANÁLISE DE IMAGEM.....	38
3	MATERIAIS E MÉTODOS	40
3.1	ANÁLISE DOS PADRÕES POLIFENOLICOS	40
3.1.1	Obtenção de Soluções padrão	40
3.1.2	Condições de análise cromatográfica	41
3.2	ANÁLISE DE <i>Schinus terebinthifolius</i>	41
3.2.1	Obtenção e tratamento da matéria prima vegetal.....	41
3.2.2	Obtenção de Solução da matéria prima vegetal.....	42
3.2.3	Obtenção do extrato bruto e fração enriquecida.....	42
3.2.4	Condições de análise cromatográfica	42
3.2.5	Cálculo do teor de ácido gálico.....	43
3.3	ANÁLISE DE <i>Anadenanthera colubrina</i>	43
3.3.1	Obtenção e tratamento da matéria prima vegetal.....	43
3.3.2	Obtenção do extrato bruto	44

3.3.3	Condições de análise cromatográfica	44
3.3.4	Cálculo do teor de catequina	44
3.4	OBTENÇÃO E PROCESSAMENTO DE IMAGEM.....	45
3.4.1	Processamento de imagem pelo software ImageJ	45
3.4.2	Processamento de imagem pelo software SidTermo	47
3.5	CONFIGURAÇÃO DE REDE NEURAL ARTIFICIAL (RNA).....	47
3.5.1	Configuração pelo GNU Octave	47
3.5.2	Configuração pelo WEKA	48
3.6	ANÁLISE QUANTITATIVA DE POLIFENÓIS POR CLAE-DAD	48
3.6.1	Condições cromatográficas	48
3.6.2	Obtenção de curva de calibração dos padrões e das amostras	50
4	RESULTADOS E DISCUSSÃO	51
4.1	ANÁLISE DE CROMATOPLACAS.....	51
4.1.1	Ácido gálico	51
4.1.2	Quercetina	52
4.1.3	Rutina	53
4.1.4	Ácido cafeico	54
4.1.5	Catequina	55
4.1.6	<i>Schinus terebinthifolius</i>	56
4.1.7	<i>Anadenanthera colubrina</i>	58
4.2	CONFIGURAÇÃO DE REDE NEURAL ARTIFICIAL (RNA).....	59
4.2.1	Ácido gálico	61
4.2.1.1	Análise de imagem por ImageJ	61
4.2.1.2	Análise de imagem por Sidtermo	66
4.2.2	Quercetina	70
4.2.2.1	Análise de imagem por ImageJ	70
4.2.2.2	Análise de imagem Sidtermo.....	74

4.2.3	Rutina	77
4.2.3.1	Análise de imagem por ImageJ	77
4.2.3.2	Análise de imagem Sidtermo.....	81
4.2.4	Ácido cafeico	85
4.2.4.1	Análise de imagem por ImageJ	85
4.2.4.2	Análise de imagem Sidtermo.....	88
4.2.5	Catequina	92
4.2.5.1	Análise de imagem ImageJ	92
4.2.5.2	Análise de imagem por <i>Sidtermo</i>	95
4.2.6	<i>Schinus terebinthifolius</i>	99
4.2.7	<i>Anadenanthera colubrina</i>	102
4.3	ANÁLISE QUANTITATIVA DE POLIFENOIS POR CLAE-DAD	105
5	CONCLUSÃO	111
	REFERÊNCIAS	111
	ANEXO A – Comprovante depósito de patente	119

1 INTRODUÇÃO

O estudo acerca de plantas medicinais apresenta relevância significativa por ampliar o conhecimento com relação ao potencial farmacológico de espécies vegetais, pelo desenvolvimento de medicamentos, por serem fontes de inúmeras moléculas bioativas que possuem estruturas ainda desconhecidas e ainda ser fundamental para a descoberta de novos protótipos de medicamentos (ZANIN et al., 2012).

No que pesem as diversas características inovadoras para quantificação de compostos bioativos, a disponibilidade de ferramentas instrumentais apropriadas à natureza das matrizes vegetais ainda é um dos principais desafios ao desenvolvimento do setor produtivo específico. O problema se torna mais crítico para espécies vegetais nativas que, embora apresentem relatos de atividades biológicas, não possuem relatos de marcadores químicos ativos ou clínicos. Portanto, a escassez de ferramentas simples e que permitam o controle contínuo da complexa composição química, ainda impede maiores avanços para o esclarecimento da eficácia e da segurança de muitas espécies medicinais brasileiras (SOARES; FERREIRA, 2017).

Assim, a técnica de impressão digital (*fingerprint*) tem encontrado maior aplicação. A técnica define um padrão único à uma espécie vegetal através da identificação de múltiplos sinais correspondentes aos constituintes químicos, de modo que, em uma amostra, os constituintes majoritários e minoritários são considerados simultaneamente. Deste modo, todos os compostos que eventualmente possam contribuir para a eficácia e/ou segurança do fitoproduto, podem ser acompanhados, auxiliando assim, no esclarecimento dos efeitos terapêuticos, os quais são frequentemente baseados nos efeitos aditivos e/ou sinérgicos dos diversos componentes (WONG et al., 2014). Portanto, o perfil químico detalhado da planta ou seu produto derivado, mesmo que qualitativo, é a principal garantia de reprodutibilidade da eficácia terapêutica.

As técnicas analíticas mais utilizadas para obtenção de impressões digitais são as cromatográficas, tais como: Cromatografia em Camada Delgada (CCD), Cromatografia Líquida de Alta Eficiência (CLAE) e Cromatografia Gasosa (CG) (DECONINCK et al., 2013).

A obtenção de *fingerprints* na análise química de drogas vegetais por CCD permanece como a mais favorável, por sua simplicidade, facilidade de acesso e pela versatilidade que permite análises de diversas classes de compostos do metabolismo

secundário através da determinação dos seus Rfs (fatores de retenção nos cromatogramas). Porém, a aplicação da CCD tem se limitado a avaliações qualitativas através da pesquisa e identificação de fitoconstituintes. Por outro lado, a técnica sofreu grande avanço na separação de misturas complexas com a introdução de placas cromatográficas com menor e melhor distribuição no tamanho das partículas, incrementado significativamente a eficiência de separação dos sistemas. O emprego das placas de CCD de alta eficiência aliada aos dispositivos automatizados (aplicadores e cubas de eluição), agregaram a reprodutibilidade que faltava a técnica, sendo assim denominada de cromatografia em camada delgada de alta eficiência (CCDAE) (MARSTON, 2011). Após a introdução da densitometria foi possível quantificar compostos nas próprias placas cromatográficas. Embora a análise quantitativa empregando densitometria seja um procedimento relativamente simples, seu emprego na rotina analítica ainda encontra alguma resistência em razão da sensibilidade, reprodutibilidade e, principalmente, do elevado custo operacional (SHEWIYO et al., 2012).

A CLAE é o método analítico usualmente utilizado para separação, identificação e quantificação de componentes em misturas complexas, sendo considerado o padrão ouro na autenticação de produtos farmacêuticos e produtos naturais devido à sua precisão, sensibilidade e reprodutibilidade. Entretanto, alguns pilares são impostos à técnica como o alto custo instrumental, tempo de análise relativamente longo e o consumo elevado de solventes (SHERMA, 2012).

Considerando as possibilidades resultantes da utilização dos agentes cromogênicos (especificidade e intensidade de coloração), os cromatogramas por CCD ou CCD-AE já tem papel consolidado como indicadores qualitativos da composição química de matérias-primas e extratos; bem como ferramenta de auxílio à identificação e autenticação destes materiais (WONG et al., 2014). Entretanto, pouco se tem explorado acerca da avaliação quantitativa destes dados cromatográficos.

Neste contexto, a exploração de imagens digitalizadas tem resultado em grandes inovações não só para avaliações de classificação, mas também de quantificação. Com a utilização de imagens digitalizadas e o auxílio de programas computacionais e técnicas de processamento de imagens, é possível obter dados quantitativos a partir de *fingerprints* cromatográficos. O procedimento é rápido e

suficientemente preciso, além de mais simples e de menor custo operacional (SANTOS, 2009).

Além disso, a obtenção de tecnologias especializadas (*hardware* e *software*) que utilizem técnicas da inteligência artificial e quimiométricas, juntamente com o processamento de imagens, tem sido amplamente difundida para a análise de imagens representadas em gráficos complexos, para avaliação de padrões de cor e diferenciação de imagens (CUSTÓDIO, 2009).

Assim, esta proposta versa sobre a aplicação de abordagem computacional para o desenvolvimento de uma ferramenta de quantificação de marcadores químicos em espécies vegetais, a partir dos dados resultantes da digitalização de placas de CCD. Para tanto, espécies reconhecidas na medicina tradicional e que possuam ao menos um marcador químico conhecido serão adotadas como modelo para desenvolvimento e validação do protocolo experimental.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

Desenvolver e avaliar uma ferramenta computacional para quantificação de polifenóis a partir de imagens digitais de “*fingerprints*” de drogas vegetais por cromatografia em camada delgada.

1.1.2 OBJETIVOS ESPECÍFICOS

- Desenvolver e otimizar as condições cromatográficas para análise de padrões de polifenóis gerando atributos de imagem para obter uma RNA;
- Desenvolver e otimizar condições cromatográficas para análise de *Schinus terebinthifolius* e *Anadenanthera colubrina*, gerando atributos de imagem para obter uma RNA;
- Desenvolver e avaliar procedimento de quantificação por CLAE-DAD;
- Desenvolver e avaliar uma ferramenta computacional que permita a quantificação de imagem(ns) correspondente(s) à(s) banda(s) cromatográfica(s) de(os) marcadore(s) químico(s);

- Avaliar e correlacionar os dados gerados pelo programa computacional a partir das imagens digitalizadas das bandas cromatográficas, com os dados quantitativos obtidos por CLAE para os padrões de polifenóis e para o marcador químico de *S. terebinthifolius* e *A. colubrina*.
- Desenvolver os *fingerprints* por CCD para os polifenóis em matrizes vegetais;
- Avaliar metodologias por CLAE para a quantificação de polifenóis na matéria prima e e/ou extratos das espécies vegetais;
- Otimizar ferramenta computacional para análise de(os) marcadore(s) químico(s) a partir da análise de *fingerprints* da matéria prima, extratos e derivados das espécies vegetais;
- Avaliar a correlação entre os dados gerados pelo programa computacional a partir das imagens digitalizadas das bandas cromatográficas do *fingerprint* por CCD, com os dados quantitativos obtidos por CLAE para os padrões de polifenóis presentes nas matrizes vegetais;

2 FUNDAMENTAÇÃO TEÓRICA

2.1 POLIFENOIS

Os polifenóis, também denominados de compostos fenólicos, são substâncias que possuem um anel aromático com uma ou mais hidroxilas, incluindo seus grupos funcionais. Eles estão amplamente distribuídos no reino vegetal, englobando desde moléculas simples a mais complexas com elevado grau de polimerização (ROCHA et al., 2011).

O elevado número de artigos científicos publicados na última década revelou o interesse da produção científica acerca dos compostos fenólicos devido às inúmeras atividades biológicas que apresentam (KEATING; MARTEL, 2018). Os polifenóis presentes nas espécies vegetais estão relacionados com a proteção, conferindo alta resistência a micro-organismos e pragas. Nos alimentos, estes compostos podem influenciar o valor nutricional e a qualidade sensorial, conferindo atributos como cor, textura, amargor e adstringência (EVERETTE et al., 2010).

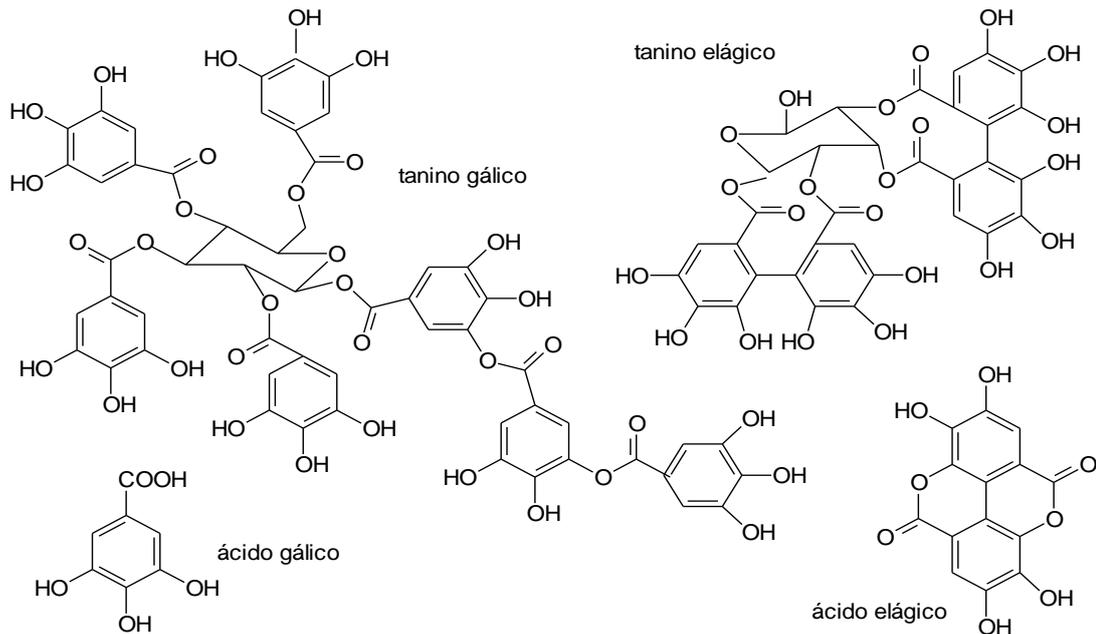
Com relação às propriedades farmacológicas, os polifenóis conferem proteção frente a algumas doenças crônicas tais como diabetes tipo 2, doenças

cardiovasculares (WILLIAMSON, 2017), obesidade (JAFFEE, 2013) e câncer (LOSADA-ECHEBERRIA, 2017; RAJAGOPAL, 2018). Adicionalmente, os compostos fenólicos também podem interagir com a microbiota intestinal modulando a composição da microbiota com um impacto incisivo na saúde humana (MARCHESI, 2016). Estudos realizados por Ghasemzadeh (2011) apontam que os polifenóis funcionam como antioxidantes naturais, sendo mais efetivos que vitamina C, E e carotenoides; sendo conhecidos por atuarem como antioxidantes, devido a capacidade de doar hidrogênio ou elétrons, inibindo assim a ação de espécies reativas de oxigênio (ROS) e reações de peroxidação lipídica (KEATING; MARTEL, 2018).

A depender da sua estrutura química os compostos fenólicos podem ser classificados como taninos, flavonoides, ácidos fenólicos e cumarinas (KEATING; MARTEL, 2018). Os taninos destacam-se devido sua ampla distribuição e por desencadarem diversas propriedades biológicas. São considerados compostos de alto peso molecular, subdivididos em hidrolisáveis e condensados (SIMÕES et al., 2016). Os taninos hidrolisáveis (Figura 1) consistem de ésteres de ácido gálico e ácido elágico glicosilados, formados a partir do ácido chiquímico, onde os grupos hidroxila do açúcar são esterificados com os monômeros destes ácidos (MONTEIRO, 2005).

O ácido gálico é um composto fenólico presente em inúmeras espécies medicinais, tais como *Anacardium occidentale* (LEITTE et al., 2016), *Libidibia ferrea* (FERREIRA et al., 2016), *Schinus terebinthifolius* (ROCHA, et al., 2019) e *Anadenanthera colubrina* (LIMA et al., 2014), as quais se destacam por apresentar propriedades de interesse terapêutico. Esse composto é amplamente utilizado como suplemento ou aditivo dietético apresentando relevância na clínica médica. Ao longo das últimas décadas, publicações relatam propriedades antivirais (WU et al., 2017), antioxidantes, anti-apoptóticas, cardioprotetoras, neuroprotetoras e anticancerígenas (KOSURU et al., 2017).

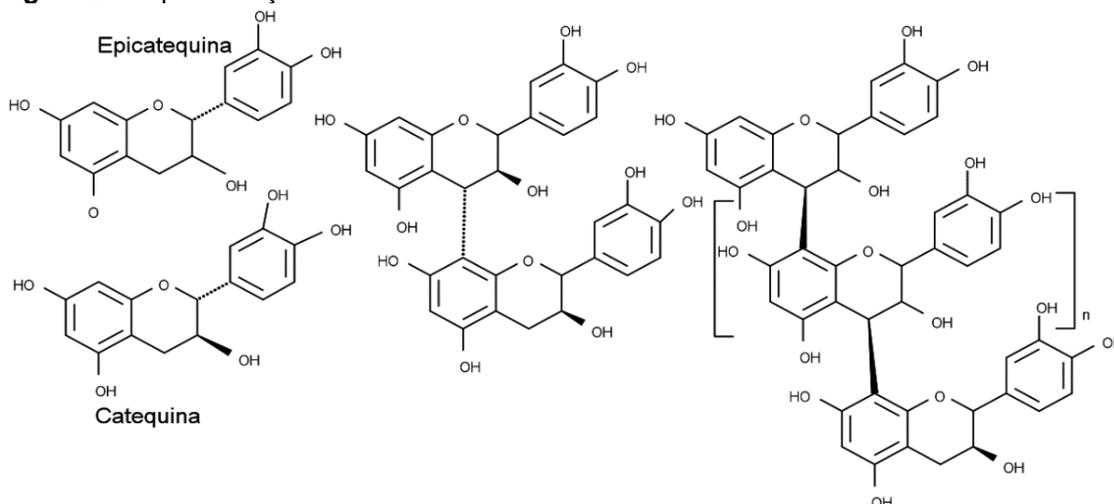
Figura 1 - Representação da estrutura dos taninos hidrolisáveis.



Fonte: BEZERRA, 2016.

Enquanto, os taninos condensados ou proantocianidinas são polímeros de flavan-3-ol e/ou flavan3,4-diol (Figura 2), resultantes do metabolismo do fenilpropanol (HASLAM, 2007), entre os principais representantes da classe está a catequina. Esse composto possui ação direta nas desordens estéticas como o envelhecimento cutâneo intrínseco e extrínseco devido à ação dos radicais livres, funciona como potente antioxidante através da doação de elétrons inibindo os danos causados ao DNA pelas espécies reativas do oxigênio (EROs), gerando um radical flavínico bem menos reativo (SENGER et al., 2010). A propriedade antioxidante está relacionada à estrutura química das catequinas, sendo potencializada a ação devido à presença de hidroxilas e radicais ligados aos anéis (YANG et al., 2010). Além dessa propriedade, outros estudos apontam que as catequinas apresentam efeitos anti-inflamatórios devido a pré inibição das enzimas ciclooxigenase 2 (COX-2) e lipoxigenase do metabolismo do ácido araquidônico, que são responsáveis por desencadear eventos inflamatórios (SENGER et al., 2010).

Figura 2 - Representação da estrutura dos taninos condensados.



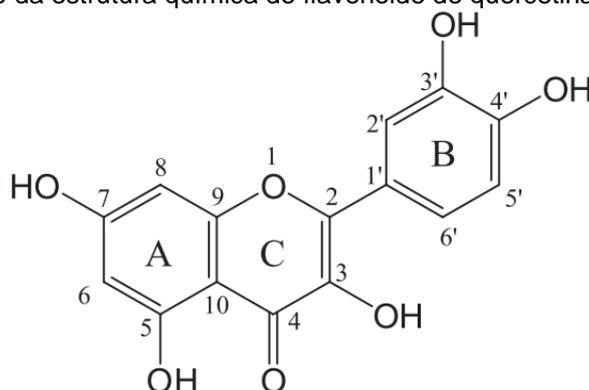
Fonte: BEZERRA, 2016.

Os taninos possuem a propriedade de formarem complexos insolúveis na presença de proteínas; supõe-se que a aplicação mais antiga desta propriedade química de taninos na tecnologia é a estabilização de proteínas de peles de animais contra a putrefação (COVINGTON, 2009).

Outra classe de compostos fenólicos com destaque para o ramo farmacêutico são os flavonoides. Esses compostos possuem estrutura de carbono $C_6-C_3-C_6$ que consistem em três anéis aromáticos: dois anéis fenilo (A e B) e um anel heterocíclico (C) (XIE et al., 2017). De acordo com o grau de hidrogenação do anel C e o local de conexão do anel B, os flavonoides podem ser divididos em várias subclasses, como flavanonas, flavonas, flavonóis, flavanonois e isoflavonoides (XIAO, 2013).

A literatura relata que os flavonoides são capazes de inativar a bomba de efluxo, desestabilizar a membrana citoplasmática e inibir β -lactamases e topoisomerase, e assim prevenir o desenvolvimento de resistência bacteriana à antibióticos (DAGLIA, 2012). A quercetina (Figura 3) é um marcador flavonoídico que merece destaque por ser amplamente encontrado em espécies medicinais. Pesquisas com relação a esse composto apontam propriedades potencialmente anticancerígenas devido aos seus comprovados efeitos antioxidantes, anti-inflamatórios e antiproliferativos que estão diretamente relacionados à sua estrutura química. Estudos experimentais disponíveis indicam que a quercetina pode modular vários RNAs mensageiros relevantes ao câncer inibindo, assim, o início e o desenvolvimento dessa patologia (KIM et al., 2019).

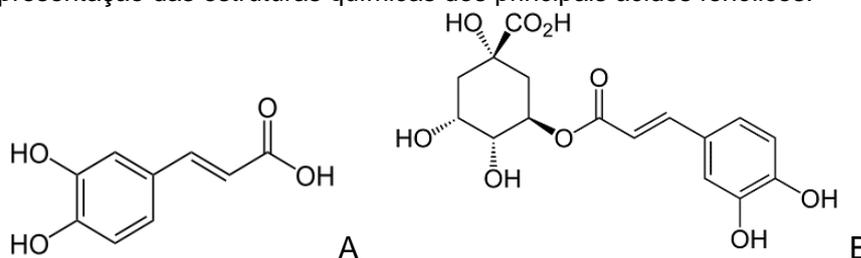
Figura 3 - Representação da estrutura química do flavonoide de quercetina.



Fonte: SIMÕES et al., 2016.

Entre os principais compostos fenólicos ácidos, estão o ácido cafeico e seu éster ácido clorogênico (Figura 4), cujas atividades biológicas são comumente associadas a propriedades antioxidantes e hepatoprotetoras (MONDAL, 2017). Devido às atividades atribuídas a esses compostos, a literatura relata a necessidade de ampliar os estudos para o desenvolvimento de medicamentos fitoterápicos (TISTAERT; DEJAEGHER; HEYDEN, 2011).

Figura 4 – Representação das estruturas químicas dos principais ácidos fenólicos.



A: Ácido Cafeico; B: Ácido Clorogênico.

Fonte: Autoria própria.

2.2 *Schinus terebinthifolius*

Schinus terebinthifolius Raddi é uma espécie medicinal nativa brasileira, pertencente à família Anacardiaceae, popularmente conhecida como aroeira, aroeira-branca, aroeira-da-praia, aroeira-do-campo, aroeira-do-brejo, aroeira-mansa, aroeira-pimenteira, aroeira-vermelha, aroeira-precoce, bálsamo, cabuí, cambuí e pimenta brasileira (SILVA et al, 2017; ROCHA, 2019). A aroeira é uma árvore de porte médio que apresenta cerca de 5 a 10 metros de altura, de copa larga e tronco revestido de casca grossa, encontrada e cultivada no Brasil. Está localizada na mata atlântica do Rio Grande do norte até o Rio Grande do Sul (SOUZA, 2011).

Levantamentos sobre o uso popular da aroeira realizado em diversas regiões brasileiras, apontaram o uso da casca de aroeira como cicatrizante de feridas, tratamento de gripe, bronquite, contra inflamações do útero, dor de dente, no tratamento de úlceras e doenças do trato urinário (BOSCOLO; VALLE, 2008; LORENZI; MATOS, 2008; GUERRA et. al, 2010; SOUZA, 2011).

O uso popular desta espécie é tão difundido que o sistema público de saúde brasileiro incluiu *S. terebinthifolius* na Relação de Plantas Medicinais de interesse ao Sistema Único de Saúde (RENISUS) (BRASIL, 2009). Essa relação é composta por espécies vegetais com potencial de avançar nos estudos acerca de produção de produtos de interesse ao SUS.

A análise fitoquímica revela a presença de taninos, flavonoides e ácidos triterpênicos nas cascas (CARLINI et al., 2010). Estudos com o emprego de cromatográfica líquida ultra-rápida acoplada a detector de arranjos de diodos (DAD) revelou que os extratos de *S. terebinthifolius* possuem um conjunto de diferentes constituintes bioativos, sendo os compostos fenólicos os principais constituintes dos extratos. Os fitoconstituintes identificados na espécie são: agliconas miricetrina, quercetina, kampferol, hexosídeo tri- *O*-gálico, hexosídeo tetra- *O*-gálico, hexosídeo hexa-*O*-gálico, ácido 4-*O*-metil-gálico, penta-*O*-galloil hexosídeo e ácido gálico (MARCH et al., 2006; LIAO et al., 2013; ABU-REIDAH et al., 2015; KUMAR et al., 2015; ERSAN et al., 2016; MUCCILLI et al., 2017; ROCHA, 2019).

A literatura relata que essa espécie apresenta diferentes atividades biológicas devido à presença de polifenóis. Rocha e colaboradores (2019) realizaram estudos com camundongos analisando vários parâmetros antropométricos e bioquímicos com a finalidade de avaliar a atividade antidiabética *in vivo* do extrato metanólico de folhas de *S. terebinthifolius*. O valor glicêmico apontou que em camundongos diabéticos os resultados foram mais efetivos quando o tratamento utilizado com extratos das folhas de *S. terebinthifolius* foi empregado do que aqueles desencadeados pelo medicamento referência metformina nas concentrações testadas.

A espécie também apresentou atividade antiulcerogênica quando extratos das cascas foram avaliados em ratos. O efeito antiúlcera foi detectado pelo aumento no volume e pH do suco gástrico (CARLINI et al., 2010). Revisões da literatura de espécies medicinais que apontam atividade antiulcerogênica relatam que essas ações estão atribuídas aos efeitos antioxidantes e protetores dos taninos (FALCÃO et al., 2008; CARLINI et al, 2010).

Outros estudos realizados apontaram também como principais propriedades farmacológicas as atividades antimicrobiana (SILVA et al., 2017a), anticancerígena (SILVA et al., 2017b), anti-hipertensiva (GLÓRIA et al., 2017), anti-histamínica (NUNES-NETO et al., 2017) e anti-inflamatória (SILVA et al., 2017b) e antioxidante (ROCHA et al., 2017).

Diante da importância biológica atribuída à espécie, é necessário o desenvolvimento de novas ferramentas que possibilitem o estudo quantitativo de marcadores da espécie, como o ácido gálico, com a finalidade de garantir maior eficácia e segurança para o desenvolvimento de produtos derivados.

2.3 *Anadenanthera colubrina*

Anadenanthera colubrina é uma espécie pertencente à família Leguminosae com crescimento em áreas de savanas e floresta tropical seca. No Brasil, onde a espécie é popularmente conhecida como angico, ocorre crescimento predominantemente em florestas sazonais e galerias ribeirinhas (MOTA et al., 2017). A espécie apresenta vasta aplicabilidade, tais como utilização para madeira, carvão e lenha, bem como na medicina popular. A madeira tem alta densidade, resistência e superfície lustrosa com alto valor decorativo, sendo utilizada para pavimentação e construção naval (LORENZI, 2009).

Estudos multidisciplinares apontam que *A. colubrina* é uma das espécies mais amplamente utilizadas na medicina popular sendo empregada para tratar problemas respiratórios, inflamação, diarreia, tosse, bronquite, gripe e dor de dente (LIMA et al., 2014). A casca é a parte mais comumente utilizada da planta, cuja disponibilidade não é limitada por variações nas condições climáticas (SARTORI et al., 2014).

A casca desta espécie é reportada, através dos estudos farmacológicos, realizados por Santos e colaboradores (2013), como agente anti-inflamatório com atividade antinociceptiva periférica, indicando a existência de compostos com potencial para o desenvolvimento de produtos farmacêuticos para o tratamento da inflamação e da dor.

Outro estudo realizado aponta que os extratos obtidos das cascas da espécie apresentam potencial antifúngico, pois o extrato e sua fração ativa inibiram o crescimento de *Candida albicans*, demonstrando forte atividade antimicrobiana. O

extrato também demonstrou natureza fungistática, sugerindo e a inibição do crescimento do microrganismo sem causar a morte (LIMA et al., 2014).

As atividades farmacológicas desta espécie estão relacionadas com a abundância em compostos fenólicos. Ao realizar a triagem fitoquímica, autores confirmaram a presença de catequinas, flavonoides, saponinas, esteroides, taninos, triterpenos e xantonas (SANTOS, 2013). Os estudos realizados por Lima e colaboradores (2014) indicaram que os polifenóis representaram a maior composição química com cerca de 53,18% do extrato. Dessa forma, diante das aplicabilidades e das atividades farmacológicas atribuídas a espécie é necessário o desenvolvimento de novas ferramentas analíticas que possibilitem a quantificação de constituintes químicos da espécie.

2.4 TÉCNICAS ANALÍTICAS DE CONTROLE DE QUALIDADE

A avaliação de drogas vegetais considerando a identificação e quantificação de marcadores químicos depende do emprego de técnicas analíticas que permitam detectar com precisão e sensibilidade a presença de fitoconstituintes. Dessa forma, o desenvolvimento de perfis químicos típicos (*fingerprints*) pode ser muito mais efetivo no auxílio à identificação/autenticação de materiais vegetais, inclusive para a análise de produtos derivados tal como extratos. A técnica de impressão digital tem o potencial de identificar a espécie, quantificar compostos ativos e detectar impurezas ou contaminantes, como herbicidas (TISTAERT; DEJAEGHER; HEYDEN, 2011). É nesse contexto que a técnica de impressão digital tem encontrado maior aplicação.

A abordagem é recomendada pela Organização Mundial de Saúde (OMS) e adotada por outras autoridades internacionais tais como *Chinese State of Food and Drug Administration* (CFDA), *US Food and Drug Administration* (FDA) e *European Medicines Agency* (EMA) (TISTAERT; DEJAEGHER; HEYDEN, 2011; SUN; CHEN, 2012).

Para obtenção de perfil químico de espécies vegetais a cromatografia líquida de alta eficiência (CLAE) merece destaque por fornecer simultaneamente informações qualitativas e quantitativas acerca dos fitocompostos. Entretanto, a técnica torna-se dispendiosa pelo emprego de elevado consumo de solventes, necessidade de detector universal, operações complexas de preparação e tratamento de amostra, e, a necessidade de recursos humanos especializados (LOESCHER et al., 2014). Além

disso, as colunas cromatográficas podem apresentar sensibilidade às variações de pH, a depender da fase móvel empregada (TISTAERT; DEJAEGHER; HEYDEN, 2011).

Outra técnica comumente utilizada para obtenção de *fingerprints* é a cromatografia gasosa (CG) por ser capaz de separar com sucesso diferentes constituintes do extrato. Entretanto, limita-se ao uso de compostos voláteis tais como óleos essenciais em preparações à base de plantas (HEYDEN, 2008).

A cromatografia em camada delgada (CCD) também é uma técnica que se destaca para obtenção de impressões digitais, sendo empregada para análise qualitativa de produtos naturais e derivados e, comumente usada para separação de componentes como etapa inicial (SALMON et al., 2012). Com os avanços da tecnologia nos últimos anos, a CCD tornou-se mais sofisticada, com o advento da cromatografia em camada delgada de alta eficiência (CCD-AE), a qual tem o potencial de fornecer informações tanto qualitativa quanto quantitativa dos constituintes químicos vegetais (HOUGHTON, 2009). Porém, a técnica ainda apresenta grande resistência do setor produtivo específico devido ao alto custo instrumental. Dessa forma, modelos estatísticos sofisticados ganham espaço para análise de fitocompostos.

2.5 QUIMIOMETRIA

A quimiometria é um ramo da ciência que aplica dados matemáticos e estatísticos para extração de informações úteis de fenômenos físicos e químicos envolvidos em um processo de fabricação (SINGH, 2013). Em decorrência da crescente sofisticação das técnicas instrumentais de análise química, impulsionada pela maior aplicabilidade de microprocessadores e microcomputadores em laboratórios, técnicas de tratamentos de dados mais complexas do ponto de vista matemático e estatístico tornaram-se necessárias para solução de problemas (DACCOL; DANTAS; POPPI, 2018; PEREIRA, 2014, 2018).

A aplicação da quimiometria é encontrada em coleta e análise de dados multivariados, calibração, modelagem de processos, reconhecimento e classificação de padrões e controle estatístico de processos, com o objetivo de alcançar o máximo de precisão e robustez (SINGH, 2013). Especificamente, no ramo farmacêutico a quimiometria encontra-se aplicada ao controle de qualidade de produtos

farmacêuticos, interpretação de testes laboratoriais no monitoramento de doenças, como ferramenta para descoberta de novos fármacos e tecnologia de bioinformática (MOKAC, 2012).

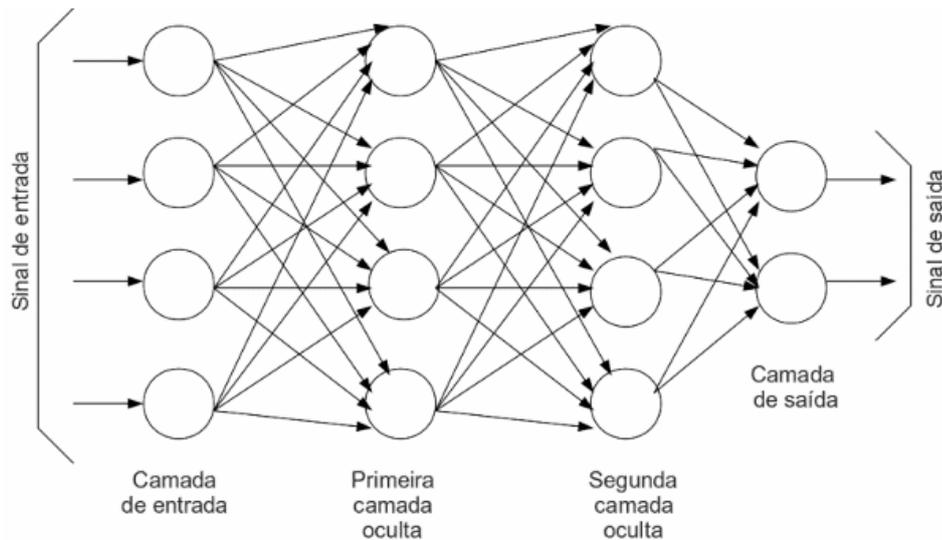
A técnica quimiométrica fundamenta-se na análise de dados, sendo considerada essencial para o resultado de todo experimento. As análises de modelos estatísticos de método dividem-se em univariada e multivariada, sendo considerada univariada quando somente uma variável é medida sistematicamente para várias amostras, o que tornou seu uso limitado nos últimos anos. Enquanto, o modelo multivariado considera a correlação entre muitas variáveis analisadas simultaneamente, permitindo maior extração de informações de um conjunto de dados (PEREIRA, 2003).

A implementação de técnicas quimiométricas com vista a garantir o controle global de um processo implica o uso de técnicas analíticas capazes de fornecer resultados precisos de maneira simples e rápida. Uma das técnicas quimiométricas empregadas para investigar a solução de problemas empregando modelos matemáticos é a rede neural artificial (RNA).

2.5.1 Rede Neural Artificial (RNA)

A RNA é um modelo inspirado no funcionamento de uma rede neural biológica, que simula o comportamento humano no processamento de informações. A vantagem da RNA sobre as técnicas estatísticas tradicionais é que o modelo não precisa ser explicitamente definido antes do início dos experimentos, no entanto é necessário conhecimento prévio das relações entre os fatores determinados no teste (KUSTRIN, 2013).

Uma RNA é formada por centenas de unidades individuais, neurônios artificiais ou elementos de processamento conectados com coeficientes denominados de pesos, que constituem a estrutura neural e são organizados em camadas. Cada elemento de processamento apresenta entradas ponderadas (camada de entrada), função de transferência (camada oculta) e uma saída (camada de saída), conforme apresentado na figura 5.



Fonte: MAZI, 2008.

O neurônio artificial é a base para o funcionamento da RNA. Sua estrutura funciona como processador que realiza uma soma ponderada dos sinais de entrada que são representadas por pesos. Estes parâmetros de rede são ajustáveis, cujo ajuste é realizado por um processo chamado treinamento ou aprendizado, sendo responsável pela extração das características dos dados e armazenamento de conhecimento das redes (BRAGA et al., 2007). Os pesos de conexão entre os neurônios são ajustados até obter um erro de previsão mínimo, com a finalidade de aproximar o resultado experimental ao máximo do real. A ativação de um sinal de saída depende do valor dessa soma de pesos e da função de ativação ou transferência, representando a parte não linear de cada neurônio. O comportamento de uma rede neural é determinado pelas funções de transferência de seus neurônios que aplicam uma determinada função matemática aos dados, gerando uma única resposta (BINOTI et al., 2014).

Atualmente, muitas formas de RNA são propostas e usadas com variações, principalmente no número de neurônios por camada, tipo de função de ativação dos neurônios, número de camadas ocultas e tipo de conexão entre os neurônios (BINOTI et al., 2014), com a finalidade de estimar o erro mínimo global.

2.5.1.1 Máquina de Aprendizado Extremo (Extreme Learning Machine - ELM)

Dentre os diversos tipos de redes neurais existentes, a Máquina de Aprendizado Extremo (Extreme Learning Machine - ELM) se destaca devido a

velocidade de treinamento e predição de dados quando comparada a outros métodos não lineares (HUANG et al., 2012). A rede ELM normalmente apresenta camada oculta única baseada em um método analítico para estimar os pesos de saída de rede, em qualquer inicialização aleatória de pesos de entrada (HUANG, 2011; CAMBRIA et al., 2013). O processo de aprendizado é realizado em lote, onde todos os dados são apresentados à rede antes do ajuste de rede (HUANG, 2011). A principal vantagem de rede ELM consiste em a camada oculta não requerer ajuste (LOUZADA, 2015). Essa ferramenta tem sido amplamente utilizada em muitos campos, como análise de sinal biomédico e análise de dados de grande porte (SINGH, 2018), devido ao baixo custo operacional.

As máquinas de aprendizado extremo são baseadas no sistema de operação de computador chamado de Kernel que além de apresentar rápido treinamento mostra desempenho de classificação e regressão de dados robustos. O núcleo Kernel permite que os processadores e dispositivos possam realizar sua função, comunicando processos e sistemas (HUANG, 2012).

Existem modelos de Kernel que são aplicados para modelagem de dados e análise de resultados, dentre os quais podem ser empregados: sigmoide, linear, polinomial, função de base radial (RBF) dilatação e erosão. O modelo de Kernel linear fundamenta-se em fazer correlações proporcionais de variáveis de dados, enquanto os demais modelos estão fundamentados na morfologia matemática não linear amplamente utilizado em processamento de imagens digitais (SANTOS, 2012).

Dessa forma, torna-se possível a construção de vários operadores projetados para aplicações específicas, como filtragem e extração de características de imagens (SOILE, 2004; HUANG, 2012). Kernel sigmoide consiste em avaliar o conjunto de dados através de uma função par e não linear como métrica de avaliação minimizando os erros de filtragem (SILVA, 2017). O modelo de dilatação aplicado para fazer análises resulta na transformação da imagem original, onde haverá o aumento de áreas mais brilhantes e a eliminação das áreas mais escuras. Enquanto o modelo de erosão consiste no aumento das áreas mais escuras e na eliminação das áreas mais brilhantes (AZEVEDO, 2015).

2.6 ANÁLISE DE IMAGEM

Nos últimos anos foram observados grandes avanços tecnológicos na área da computação científica. Um dos campos em que a computação científica fez incursões particulares tem sido na área de imagens. O sistema de análise de imagens objetiva implantar um método que possibilite fácil acesso e baixo custo operacional. Os estudos acerca de análises de imagens requerem uma variedade de diferentes softwares para adquirir e analisar dados, sendo crucial a conectividade entre essas ferramentas (SCHNEIDER; RASBAND; ELICEIRI, 2012).

A aplicação de dispositivos de digitalização, a captura e processamento de imagens associados a técnicas cromatográficas e quimiométricas está se tornando um campo atraente e promissor da ciência de separação (SCHNEIDER; RASBAND; ELICEIRI, 2012; ABOU-DONEA, 2014; GULZEMERIC, 2017). A análise multivariada de imagens é essencial para aquisição adequada de dados provenientes de imagens. Nesse sentido, variáveis obtidas como intensidade de cinza de pixels, área de pico e valores médios de cinza são utilizados como dados de entrada para o desenvolvimento de uma rede neural artificial (GULZEMERIC, 2017).

A análise de imagem é fundamentada em dados matemáticos considerando as operações básicas de comparação, subtração, intersecção e união, tais operações só podem ser aplicadas a imagens que respeitem a relação de ordem entre os pixels. Apenas imagens em nível de cinza e binárias mantêm a relação de ordem (SOILE, 2004; SANTOS, 2012).

A literatura aponta alguns programas de processamento de imagens empregados para medições quantitativas associados a técnicas quimiométricas: ImageJ, JusTLC, Sorbfil e Sidtermo. Análises realizadas por Schneider (2012) e Abou-Donia (2014) apontam que o *software* ImageJ apresentou melhor desempenho por possuir maior sensibilidade, linearidade e precisão para os problemas propostos. O *software* SidTermo tem como vantagem a não interferência do operador no processamento de imagens, o que resulta em maior predição de resultados e aproximação dos resultados experimentais ao máximo do real (SANTANA et al., 2018).

Os processamentos de tais ferramentas de imagem baseiam-se na extração de características e reconhecimento de objetos. O reconhecimento é fundamentado em descrições de formas, cujas características escalares são, por exemplo, área,

dimensão e descritores de Fourier (OLIVEIRA et al., 2012). No *software* ImageJ medidas de imagem são selecionadas em pixels para gerar atributos de imagens, que correspondem a dados matemáticos representativos da imagem em estudo. Dentre medidas de imagem que podem ser selecionadas na ferramenta de imagem são: área, que corresponde à área da imagem de interesse; valor médio de cinza, cujo valor corresponde à soma dos valores de cinza de todos os pixels na seleção dividida pelo número de pixels; desvio padrão dos valores de cinza para gerar o valor médio de cinza; valor de cinza modal, que corresponde o valor de cinza que mais aparece na área de seleção; valores mínimos e máximos de cinza; perímetro, o comprimento de limite externo da seleção; elipse, usa os títulos maior e menor ângulo da seleção; circularidade, o valor quanto mais próximo de 1 indica um círculo perfeito e quanto mais próximo de 0 indica um polígono cada vez mais alongado; diâmetro de Ferret, que corresponde a distância mais longa entre dois pontos ao longo do limite de seleção; densidade integrada, cujo valor é equivalente ao produto de área e valor de cinza médio e (SCHNEIDER et al., 2012).

No *software* Sidtermo as características de textura de uma imagem podem ser extraídas através do conceito de matrizes de co-ocorrência, descrito por Haralick. Este modelo consiste em uma contagem de diferentes níveis de cinza que ocorrem em uma imagem. A matriz de co-ocorrência é uma matriz quadrado, cujo tamanho corresponde a quantidade de níveis de cinza da imagem a ser analisada (OLIVEIRA, 2012). As combinações de ocorrência entre os níveis de cinza são calculadas nos ângulos 0,45, 90 e 135, os outros ângulos são calculados via simetria. Após os cálculos dessa matriz outra é calculada, essa nova matriz é a probabilidade de ocorrência das combinações entre os níveis de cinza. Com essa matriz de probabilidade de ocorrência são feitos cálculos dos atributos de textura. Os atributos de textura mais empregados são a energia, entropia, variância (contraste), homogeneidade, dissimilaridade e medidas de correlação. A energia permite calcular o grau de homogeneidade das medidas de distância; a entropia mede o nível de desordem das distâncias. Para o cálculo dos recursos de textura Haralick quatro sentidos da adjacência são definidos, as estatísticas são calculadas para matrizes de co-ocorrência geradas utilizando cada uma dessas direções de adjacência como apresentado na figura 6.

Figura 6 - Quatro sentidos da vizinhança definido para o cálculo de os recursos de textura Haralick.



Fonte: http://murphylab.web.cmu.edu/publications/boland/boland_node26.html

Os momentos de Zernick representam outra forma de extrair características de imagens utilizando quantidade reduzida de descritores. Os polinômios de Zernick representam um modelo matemático preciso que captura o formato global, apresenta sensibilidade a ruídos e preserva informações suficientes. As características importantes desse modelo são a invariância a rotação, translação e escala (JOAQUIM, 2005).

O estudo utilizando ferramentas de imagem no âmbito farmacêutico apresentou êxito para investigação de radicais livres (OLECH, 2012), determinação de catecolaminas (SIMA; CASONI; SÂRBU, 2013), análise de compostos fenólicos em vinho (KUSTRIN, 2015) e avaliação da atividade inibitória da acetilcolinesterase de diferentes extratos de *Amaryllidacea* (ABOU-DONIA, 2014).

Dessa forma, o estudo de imagens vigora como uma ferramenta simples, acessível e de baixo custo instrumental, funcionando como alternativa para estudos acerca de constituintes químicos de interesse farmacêutico com a finalidade de ser adicionado ao controle de qualidade de fitoprodutos.

3 MATERIAIS E MÉTODOS

3.1 ANÁLISE DOS PADRÕES POLIFENOLICOS

3.1.1 Obtenção de Soluções padrão

Preliminarmente, foram realizadas análises com as soluções de padrões polifenolicos para posteriormente realizar testes com o material vegetal. Para as análises dos padrões em cromatografia em camada delgada, soluções padrão foram preparadas em metanol e depois diluídas apropriadamente para obtenção de cinco concentrações diferentes. As soluções padrão e respectivas concentrações estão apresentadas no quadro 1.

Quadro 1 - Concentrações e sistemas de eluição para os respectivos padrões empregados para preparação de cromatoplasmas.

Padrões	Concentração (µg/mL)	Sistemas de eluição
Ácido gálico	20,0; 22,5; 25,0; 27,5;	Tolueno:AcOEt:HCOOH:MeOH ¹
Quercetina	7,5; 10,0; 12,5; 15,0; 17,5	Tolueno:AcOEt:HCOOH ²
Rutina	15,0; 18,0; 21,0; 24,0;	AcOEt:HCOOH:H ₂ O ³
Ácido cafeico	25,0; 27,5; 30,0; 32,5;	Tolueno:AcOEt:HCOOH ⁴
Catequina	12,0;14,0;16,0;18,0;20,0	Tolueno:AcOEt:HCOOH:MeOH ¹

¹3:3:0,8:0,2; ²6:4:0,3; ³90:5:5; ⁴4,5:3:0,2.

AcOEt - Acetato de etila; HCOOH - Ácido fórmico; H₂O - Água.

Fonte: Autoria própria.

3.1.2 Condições de análise cromatográfica

As soluções padrão foram aplicadas individualmente à cromatoplasmas de gel de sílica 60 - F₂₅₄ (Macherey-Nagel®, Alemanha), com o auxílio de aplicador semiautomático Linomat® V (Camag®, Suíça) monitorado pelo software WinCats® (Camag®, Suíça). As condições de operação empregadas foram: velocidade de dispensação da seringa, 100 s.µL⁻¹; largura da banda, 3 mm; espaço entre as bandas, 7,7 mm; posição, 5 mm antes do início. Após aplicação do padrão, as placas foram desenvolvidas em câmara horizontal (20 x 10 cm) (Camag®, Suíça), previamente saturadas com o sistema de eluição apresentados no quadro 1. Por fim, as cromatoplasmas foram secas a temperatura ambiente durante 20 minutos.

No total, 50 curvas de calibração de cada solução padrão foram preparadas com a finalidade de gerar um conjunto de dados de imagem que sejam suficientes para o desenvolvimento de RNA.

3.2 ANÁLISE DE *Schinus terebinthifolius*

3.2.1 Obtenção e tratamento da matéria prima vegetal

A matéria prima vegetal, constituída das cascas de *Schinus terebinthifolius*, foi coletada em Recife-PE, e levada à identificação no Instituto Agrônomo de Pernambuco – IPA sob o número de tombamento 87539. O acesso do material foi cadastrado no Sistema Nacional de Gestão do Patrimônio Genético (Sisgen) sob

número de cadastro A29EC22. Após coleta, foi realizada estabilização através de secagem em estufa (modelo 82/480, Lucadema®) de ar circulante (40 °C) por um período de 7 dias, e depois pulverizadas em moinho industrial (tipo Willye, TE-680, Tecnal®).

3.2.2 Obtenção de Solução da matéria prima vegetal

Para análise da matéria prima vegetal, 1 g da amostra (cascas) foi pesado e colocado em balão de fundo redondo, adicionado de 10 mL de metanol. Em seguida foi aquecido sob refluxo por 10 minutos e ao final filtrado em algodão.

3.2.3 Obtenção do extrato bruto e fração enriquecida

Foi pesada uma quantidade de droga correspondente a 10 g para 100 mL de acetona: água (7:3, v/v) e deixado sob refluxo durante 30 minutos. Posteriormente, foi filtrado sob vácuo e concentrado sob pressão reduzida (RV10 Basic, IKA®) para retirada do solvente orgânico. Em seguida, congelado a -80°C durante 3 dias e liofilizado (Modelo L101, Liotop®) originando o extrato bruto (EB).

Para obtenção da fração enriquecida, 10 g do EB foram exatamente pesados e reconstituídos em 100 mL de água. Em seguida, foi adicionado à solução aquosa 100 mL de acetato de etila e particionado por 10 vezes. Ao final, as frações foram reunidas e concentradas para total eliminação do solvente orgânico, congeladas e então liofilizadas, originando a fração acetato de etila (FAE).

3.2.4 Condições de análise cromatográfica

Para desenvolvimento das cromatoplasmas foram preparadas soluções em metanol da MP, EB e FAE na concentração de 1 mg/mL. Para desenvolvimento das cromatoplasmas foram retiradas quantidades correspondentes a 2 µg/mL das soluções de EB e MP, da solução FAE foram retiradas quantidades equivalentes a 30, 32,5, 35, 37,5 e 40 µg/mL e aplicadas individualmente à cromatoplasmas de gel de sílica 60 - F₂₅₄ (Macherey-Nagel®, Alemanha), com o auxílio de aplicador semiautomático Linomat® V (Camag®, Suíça) monitorado pelo software WinCats® (Camag®, Suíça).

As condições de operação empregadas foram: velocidade de dispensação da seringa, 100 s.µL⁻¹; largura da banda, 5,0 mm; espaço entre as bandas, 8,0 mm; posição, 7,0 mm antes do início. Após aplicação das amostras, as placas foram desenvolvidas em câmara horizontal (20 x 10 cm) (Camag®, Suíça), previamente saturadas com o sistema de eluição tolueno: acetato de etila: ácido fórmico: metanol (3:3:0,8:0,2, v/v/v/v). Por fim, as cromatoplasacas foram secas a temperatura ambiente durante 20 minutos.

No total, 12 curvas de calibração de cada solução amostra foram preparadas com a finalidade de gerar um conjunto de dados de imagem para o desenvolvimento de RNA.

3.2.5 Cálculo do teor de ácido gálico

O teor de ácido gálico foi calculado pela equação:

$$\frac{C_p \times AA}{m \times A_p} \times FD \times 100$$

Onde: Cp – concentração do padrão; AA – absorvância da amostra; FD – fator de diluição; Ap – absorvância do padrão; m – massa da matéria prima/extrato bruto/frações enriquecidas.

3.3 ANÁLISE DE *Anadenanthera colubrina*

3.3.1 Obtenção e tratamento da matéria prima vegetal

A matéria prima vegetal, constituída das cascas de *Anadenanthera colubrina*, foi coletada no município de Caruaru – PE, e levada à identificação no Instituto Agrônômico de Pernambuco – IPA sob o número de registro 89757. O acesso do material foi cadastrado no Sistema Nacional de Gestão do Patrimônio Genético (Sisgen) sob número de cadastro AD4E5C1. Após coleta, foi realizada estabilização através de secagem em estufa (modelo 82/480, Lucadema®) de ar circulante (40 °C) por um período de 48 horas e posteriormente pulverizadas em moinho industrial (tipo Willye, TE-680, Tecnal®).

3.3.2 Obtenção do extrato bruto

Foram adicionados 10 g de cascas pulverizadas de *Anadenanthera colubrina* a um balão de fundo redondo de 250 mL, em seguida adicionou-se 100 mL do líquido extrator (acetona: água, 7:3). A extração ocorreu sob refluxo em banho maria a 75 °C por 30 minutos. Após resfriamento, a solução foi filtrada sob vácuo e concentrada sob pressão reduzida (RV10 Basic, IKA®) para retirada do solvente orgânico. Em seguida, foi congelada a -80 °C durante 3 dias e liofilizada (Modelo L101, Liotop®) originando o extrato bruto (EB).

3.3.3 Condições de análise cromatográfica

Para análise da solução de extrato bruto, 10 mg foram pesados e diluídos em 1 mL de metanol, obtendo-se uma solução de 10 mg/mL. Foram aplicadas quantidades de 300; 325; 350; 375 e 400 µg/mL da solução de EB com o auxílio de aplicador semiautomático Linomat® V (Camag®, Suíça) em placas de gel de sílica 60 - F₂₅₄ (Macherey-Nagel®, Alemanha). As condições de operação empregadas foram: velocidade de dispensação da seringa, 100 s.µL⁻¹; largura da banda, 5 mm; espaço entre as bandas, 9 mm; posição, 5 mm antes do início. Após aplicação as cromatoplasmas foram eluídas em câmara horizontal (20 x 10 cm) (Camag®, Suíça), previamente saturada com o sistema de eluição composto por tolueno: acetato de etila: ácido fórmico: metanol (3:3:0,8:0,2, v/v/v/v), após secagem as cromatoplasmas foram reveladas com vanilina clorídrica e fotografadas em 254 e 366 nm com auxílio de fotodocumentador (MultiDoc-It Imaging System®, Mod. 125, UVP, EUA) munido de câmera digital (Rebel T3, EOS 1100 D, Canon®, Japão).

No total, 12 curvas de calibração de cada solução amostra foram preparadas com a finalidade de gerar um conjunto de dados de imagem para o desenvolvimento de RNA.

3.3.4 Cálculo do teor de catequina

O teor de catequina foi calculado através da equação:

$$\frac{C_p \times AA}{m \times A_p} \times FD \times 100$$

Onde: C_p – concentração do padrão; AA – absorvância da amostra; FD – fator de diluição; A_p – absorvância do padrão; m – massa da matéria prima/extrato bruto/frações enriquecidas.

3.4 OBTENÇÃO E PROCESSAMENTO DE IMAGEM

A observação das cromatoplasmas e a aquisição das imagens foi realizada com o auxílio de fotodocumentador (MultiDoc-It Imaging System[®], Mod. 125, UVP, EUA) munido de câmera digital (Rebel T3, EOS 1100 D, Canon[®], Japão). Após digitalização, as imagens foram processadas com o auxílio de duas ferramentas de imagem: os *softwares* ImageJ (SCHNEIDER et al., 2012) e SidTermo (SANTANA, 2018). O programa ImageJ permite o cálculo da área e de valores estatísticos de pixels selecionados pelo usuário e suporta funções padrões de processamento de imagens como alterações no contraste, *sharpening*, suavizações, detecções de bordas etc. (HAWRYL et al., 2016). Enquanto o software SidTermo permite a análise de todo campo experimental da imagem sem interferência de ajustes do operador.

3.4.1 Processamento de imagem pelo software ImageJ

Todas as imagens foram tratadas utilizando a técnica de contraste preto no branco e os limiares foram ajustados até suavização e adequada delimitação das bandas. O método de limiar escolhido que apresentou melhor adequabilidade das imagens está apresentado no quadro 2.

Quadro 2 - Condições de imagem utilizadas no ImageJ para análise.

Limiar de imagem		Ácido gálico	Quercetina	Rutina	Ácido cafeico	Catequina
254 nm	Thresholdingmethod	Intermodes	Intermodes	Isodata	Huang	Default
	Thresholdingcolor	B&W	B&W	B&W	B&W	B&W
	Color back	LAB	LAB	LAB	HSB	B&W
366 nm	Thresholdingmethod	Default	Default	Huang	Default	Default
	Thresholdingcolor	B&W	B&W	B&W	B&W	B&W
	Color back	RGB	HSB	HSB	LAB	B&W

Fonte: Autoria própria.

As medidas de imagem selecionadas em pixels para verificar as semelhanças de imagem encontram-se apresentadas no quadro 3. Os valores das medidas foram selecionados em pixels quadrados.

Quadro 3 - Medidas de imagem selecionadas para análise.

Medidas de imagem	
Área	Perímetro
Valor de cinza médio	Elipse
Desvio padrão	Circularidade
Valor de cinza modal	Diâmetro de Ferret
Valores mínimos e máximos de cinza	Densidade integrada
Retângulo delimitador	Mediana

Fonte: Autoria própria.

A opção *display label* foi selecionada para detectar as medidas correspondentes às bandas em análise. Após obtenção dos atributos de imagem foram gerados formatos de arquivo de atributo-relação (ARFF) para estruturar a base de dados manipulada pela plataforma de classificação. Posteriormente, os arquivos

ARFFs foram processados pelo *software* GNU Octave® (versão 4.0.3), cujas condições de análise estão descritas na próxima sessão.

3.4.2 Processamento de imagem pelo software SidTermo

As imagens de cada concentração (Quadro 1) foram analisadas isoladamente. A lista de imagens de cada concentração foi processada no SidTermo, com a finalidade de gerar os atributos de imagem correspondente à banda cromatográfica e os arquivos ARFFs. Esses arquivos foram, posteriormente, processados pelo *software* GNU Octave® (versão 4.0.3), cujas condições de análise estão descritas na próxima sessão. Os atributos de configuração selecionados no SidTermo foram os momentos de Haralick e Zernick. O número de níveis discretos foi correspondente a 8 para processamento das imagens.

3.5 CONFIGURAÇÃO DE REDE NEURAL ARTIFICIAL (RNA)

3.5.1 Configuração pelo GNU Octave

Para desenvolvimento da rede neural artificial (RNA) dos padrões polifenólicos e das espécies medicinais de interesse foi utilizado o *software* GNU Octave® (versão 4.0.3) que fornece uma linha de comando conveniente para resolução de problemas lineares e não lineares numéricos (JONH et al., 2015).

Foi utilizada a Máquina de Aprendizado Extremo em várias configurações para o Kernel da camada intermediária. Para modelagem e predição dos resultados foram testadas as variáveis de Kernel: sigmoide ($n=100$); polinomial ($P=2,3,4,5$); função de base radial RBF gaussiana ($n=100$); dilatação ($n=100$) e erosão ($n=100$).

Dois métodos foram utilizados para análise das variáveis: divisão percentual e validação cruzada (10 k-folds). O método de divisão percentual consiste em dividir o conjunto total de dados em dois subconjuntos mutuamente exclusivos, um corresponde ao treinamento para estimação dos parâmetros e outro para teste. Enquanto o método por validação cruzada consiste em avaliar a capacidade de generalização de um modelo, a partir de um conjunto de dados. Das 50 curvas de calibração preparadas, 75% foram utilizadas para o treinamento da rede e 25% para teste. Foram realizadas 30 leituras dos dados de imagem para cada variável de kernel

aplicada. Após a estimação do modelo, os dados do teste e o erro de predição foram calculados e analisados.

3.5.2 Configuração pelo WEKA

Para o desenvolvimento das redes neurais artificiais (RNAs) foi utilizada a função *Multilayer Perceptron software Weka* (versão 3.8.1), o teste foi realizado através de validação cruzada empregando o método k-fold, cujo valor $k = 10$ com a finalidade de estimar a acurácia do modelo. Ao total, 30 atributos foram avaliados em 108 instâncias. O número de treinamentos foi arbitrado como 500. Os parâmetros analisados para definir a RNA com resultados mais precisos foram testados com uma camada oculta e 100 neurônios, taxa de aprendizagem do sistema foi de 0,1 e *momentum* de 0,2. As medidas de imagem representaram os dados de entrada do sistema e a concentração o dado de saída.

3.6 ANÁLISE QUANTITATIVA DE POLIFENOIS POR CLAE-DAD

3.6.1 Condições cromatográficas

As análises foram realizadas em cromatógrafo líquido acoplado com detector de arranjo de fotodiodos (DAD), desgaseificador e sistema de injeção automática (Ultimate 3000, ThermoScientific®). O sistema foi controlado pelo *software* Chromeleon®.

As condições cromatográficas para as análises foram estabelecidas a partir de ensaios preliminares utilizando como ponto de partida os seguintes parâmetros: Composição da fase móvel - A) Água purificada, B) Acetonitrila ou Metanol, C) Ácido acético, ácido fosfórico ou ácido trifluoracético; Fluxo: 0,5 – 1,5 mL/minuto; Coluna: fase reversa C₁₈. As condições do método que apresentaram separação das substâncias de interesse, simetria de picos, número de pratos teóricos e resolução estão apresentados no quadro 4.

Quadro 4 - Condições cromatográficas utilizadas nas análises por CLAE.

Parâmetros	Padrões	<i>S. terebinthifolius</i>	<i>A. colubrina</i>
Modo Gradiente	0-10 min, 10%B 10-15 min, 10-25%B 15-25 min, 25-40%B 25-30 min, 40-75%B 30-31 min, 75-10%B	0-10 min, 10-25%B 10-20 min, 25-40%B 20-25 min, 40-75%B 25-28 min, 75-10%B 28-32 min, 10%B	0-15 min, 30-50%B 15-16 min, 50-75%B 16-17 min, 75-30%B 17-18 min, 30%B
Coluna	250 mm x 4,6 mm (C ₁₈ ; 5 µm); pré-coluna (C ₁₈ , 3,9 µm)		
Fase móvel	Água acidificada com TFA 0,05% (A); Metanol acidificado com TFA 0,05% (B)		
Fluxo	0,8 mL/min		
Volume de injeção	20 µL		
Temperatura	23 ± 2 °C		
λ	254, 270 e 350 nm		

Onde: TFA: Trifluoroacético.

Fonte: Autoria própria.

3.6.2 Obtenção de curva de calibração dos padrões e das amostras

Foram preparadas curvas de calibração dos padrões de ácido gálico, quercetina, rutina, ácido cafeico e catequina nas condições cromatográficas descritas acima (Quadro 4). As soluções padrões e respectivas concentrações foram as mesmas descritas no quadro 1, para análise por cromatografia em camada delgada.

Para obtenção da curva de calibração com cinco pontos das amostras de aroeira foram preparadas soluções na faixa de concentração de 800 a 1200 µg/mL. Enquanto para análise de da espécie de *A. colubrina* foram obtidas curvas de calibração empregando a faixa de concentração de 400 a 1200 µg/mL. As curvas foram obtidas a partir da média de três curvas autênticas.

Os parâmetros estatísticos de resposta de linearidade foram estabelecidos através da análise de regressão linear (Excel®), sendo analisados os parâmetros: R^2 , resíduos, interseção/inclinação, e os limites de detecção (LD) e os limites de quantificação (LQ). As equações utilizadas para calcular os parâmetros de LQ e LD estão descritas abaixo, conforme preconizadas pela RDC 166/2017.

$$LD = 3,3 \cdot \left(\frac{S}{I}\right) \quad LQ = 10 \cdot \left(\frac{S}{I}\right)$$

Onde o S = desvio padrão do intercepto com o eixo do Y e I = inclinação da curva.

4 RESULTADOS E DISCUSSÃO

4.1 ANÁLISE DE CROMATOPLACAS

Na otimização das placas cromatográficas, os sistemas de eluição foram de suma importância para obtenção de uma imagem com qualidade para possível quantificação. As bandas cromatográficas mais delimitadas e uniformes apresentaram-se mais fidedignas as imagens sem dispersão de pontos após tratamento e digitalização. Os eluentes indicados no quadro 1, para os respectivos padrões e amostras analisadas permitiram a obtenção de bandas cromatográficas que após análise de imagem, toda região de interesse da imagem fosse considerada.

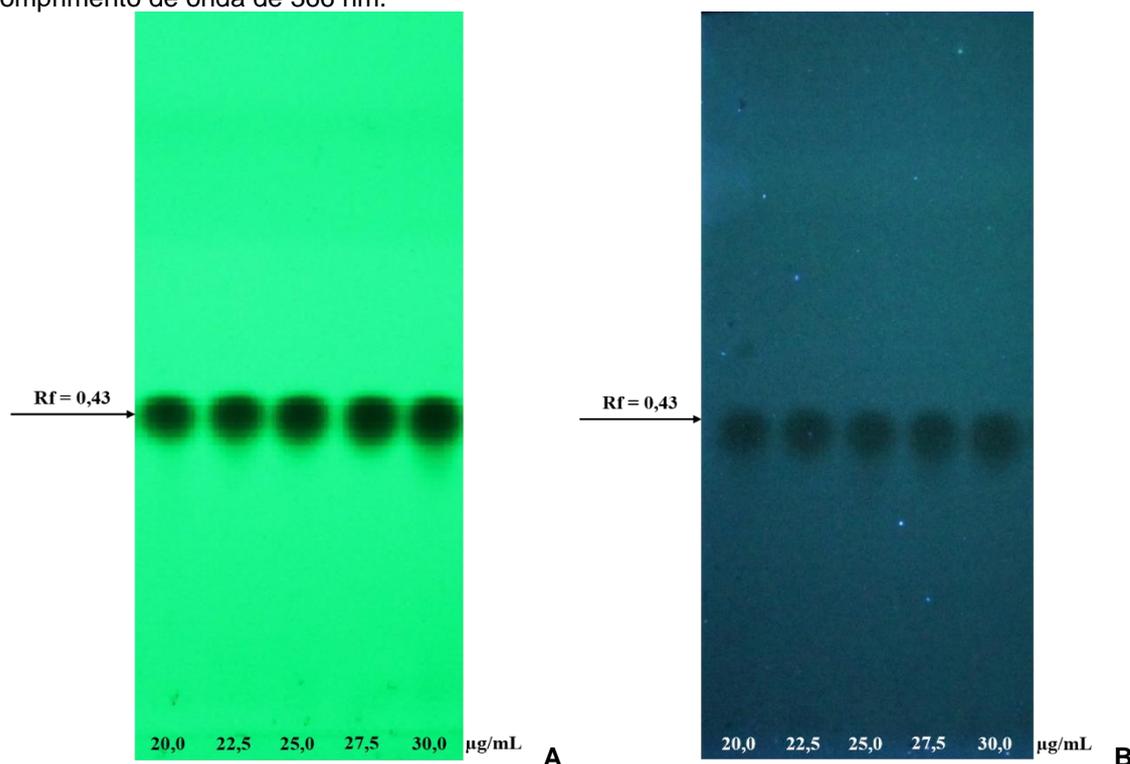
4.1.1 Ácido gálico

A análise do ácido gálico por CCD resultou na obtenção de 50 cromatoplasmas. Antes de fixar o sistema de eluição mais adequado como descrito no quadro 1, outro sistema de solventes foi avaliado com a finalidade de identificar o mais apropriado para realizar a análise das bandas cromatográficas. Inicialmente, o sistema de eluição usado acetato de etila: ácido fórmico: água na proporção 90:5:5 (v/v/v) apresentou boa eficiência de separação. No entanto, a banda não se apresentou tão delimitada, não sendo interessante para análise de imagens, pois pode ocasionar na falta de precisão de resultados. Na tentativa de encontrar uma banda mais delimitada para análise no programa de imagem, outro sistema de eluição foi utilizado tolueno: acetato de etila: ácido fórmico: metanol, na proporção 3:3:0,8:0,2 (v/v/v/v). Ao utilizar essa fase móvel foi verificado melhor eficiência de separação do composto e as bandas apresentaram-se mais delimitadas. Assim, o último sistema foi adotado como ideal para análise das imagens resultantes da observação sob 254 nm e 366 nm que estão apresentadas na figura 7.

O programa de imagem (ImageJ) apresentou sensibilidade para gerar dados de imagens característicos para cada concentração empregada. Acerca das imagens apresentadas na figura 8 é importante ressaltar que não foi necessário utilizar agentes cromogênicos em nenhuma das placas cromatográficas para visualizar as bandas, sendo possível gerar os atributos de imagem sem utilizar reagentes de derivatização, tornando o processo de quantificação de compostos bioativos menos dispendioso.

O valor de Rf correspondente à banda do padrão de ácido gálico foi de 0,43. O Rf é um parâmetro essencial na análise qualitativa empregado na cromatografia em camada delgada, pois auxilia na investigação de compostos químicos.

Figura 7 - Cromatoplasmas desenvolvidas para análise de imagem do padrão ácido gálico. Faixa de concentração 20-30 µg/mL. Sistema de eluição - tolueno: AcOEt: ácido fórmico: MeOH (3:3:0,8:0,2). A - Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B - Cromatoplasma visualizada sob o comprimento de onda de 366 nm.



Fonte: Autoria própria.

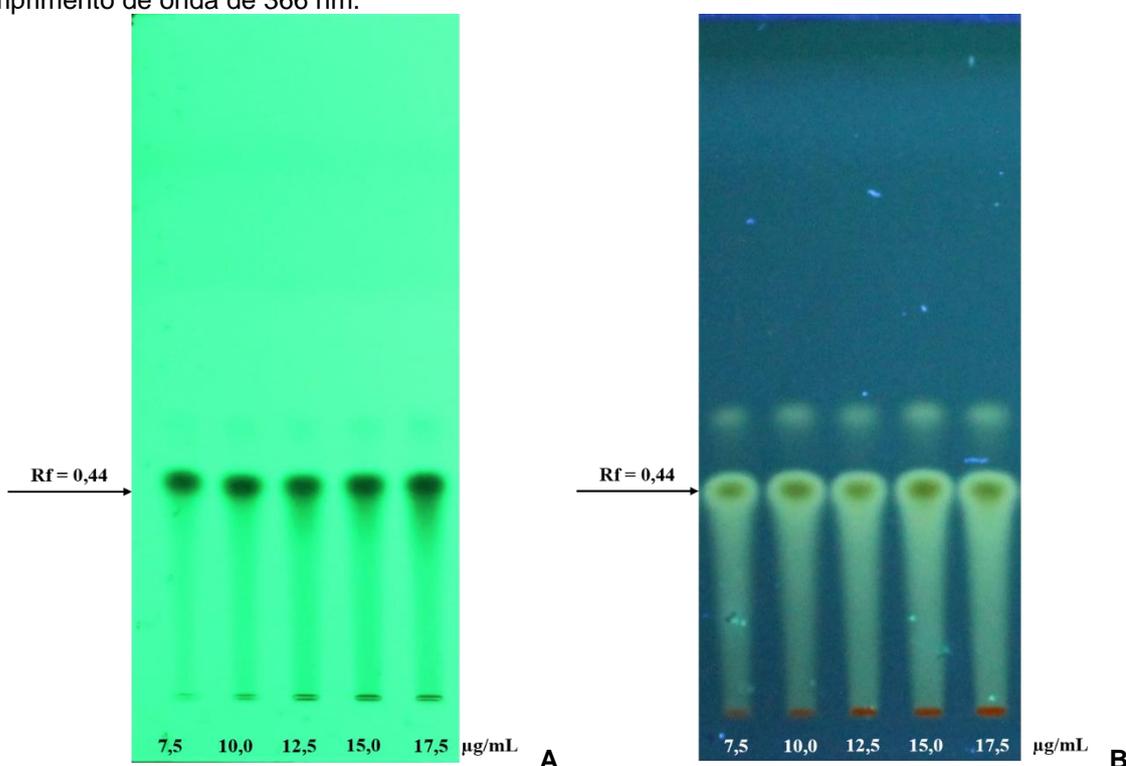
4.1.2 Quercetina

Para o padrão quercetina dois sistemas de eluição foram testados. O primeiro sistema testado foi composto pelos solventes acetato de etila: ácido fórmico: água na proporção 90:5:5 (v/v/v). A banda cromatográfica mostrou-se com presença de cauda e o valor de Rf foi igual a 0,8. Para o desenvolvimento das cromatoplasmas considerou-se como valor de Rf ideal igual ou abaixo de 0,5, pois quando for realizar a análise do material vegetal o interessante é que apareça o máximo de bandas na espécie com a finalidade de mostrar o perfil químico mais detalhado. O segundo sistema de eluição testado foi composto por tolueno: AcOEt: HCOOH (6:4:0,3), cujo valor de Rf foi igual a 0,44 como pode ser visualizado na figura 8, para as cromatoplasmas obtidas em 254 nm e 366 nm. Foi verificada a presença de cauda do padrão na placa cromatográfica,

para tentar solucionar e evitar os arrastes foram testadas concentrações menores na faixa de 2,5-5,0 μL , no entanto não foi possível detectar as bandas no ImageJ quando concentrações inferiores foram aplicadas. Isso provavelmente se justifique devido ao limite de detecção pelo programa de imagem. Dessa forma, a análise foi realizada empregando as concentrações na faixa de 7,5 – 17,5 $\mu\text{g/mL}$.

Na parte superior da cromatoplaça registrada sob o comprimento de onda de 366 nm é verificada a presença de bandas que não correspondem ao padrão de quercetina. Essas bandas correspondem possivelmente às impurezas do padrão.

Figura 8 - Cromatoplaças desenvolvidas para análise de imagem do padrão quercetina. Faixa de concentração: 7,5-17,5 $\mu\text{g/mL}$. Sistema de eluição - tolueno: AcOEt: HCOOH (6:4:0,3). A – Cromatoplaça visualizada sob o comprimento de onda de 254 nm; B – Cromatoplaça visualizada sob o comprimento de onda de 366 nm.



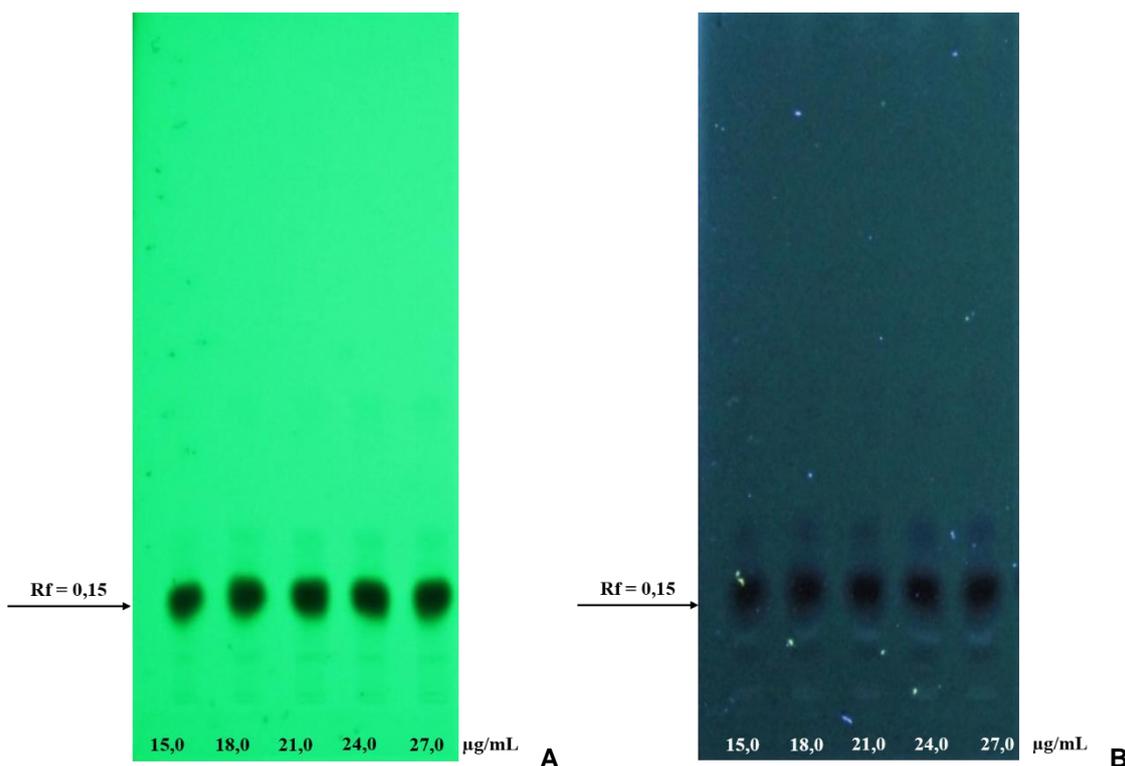
Fonte: Autoria própria.

4.1.3 Rutina

Para o padrão do flavonoide rutina foi evidenciado banda com Rf correspondente a 0,15 e as bandas mostraram-se com formas definidas, como pode ser visualizado nas imagens das cromatoplaças em 254 nm e 366 nm que estão apresentadas na figura 9. Assim como para os padrões anteriormente descritos, as ferramentas de imagem utilizadas possibilitaram identificar as variações das bandas

quanto à forma e intensidade de coloração, e assim atributos de imagem foram gerados para associar com a respectiva concentração.

Figura 9 - Cromatoplasmas desenvolvidas para análise de imagem do padrão Rutina. Faixa de concentração: 15-27 $\mu\text{g/mL}$. Sistema de eluição - AcOEt: HCOOH: H₂O (90:5:5). A – Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B - Cromatoplasma visualizada sob o comprimento de onda de 366 nm.

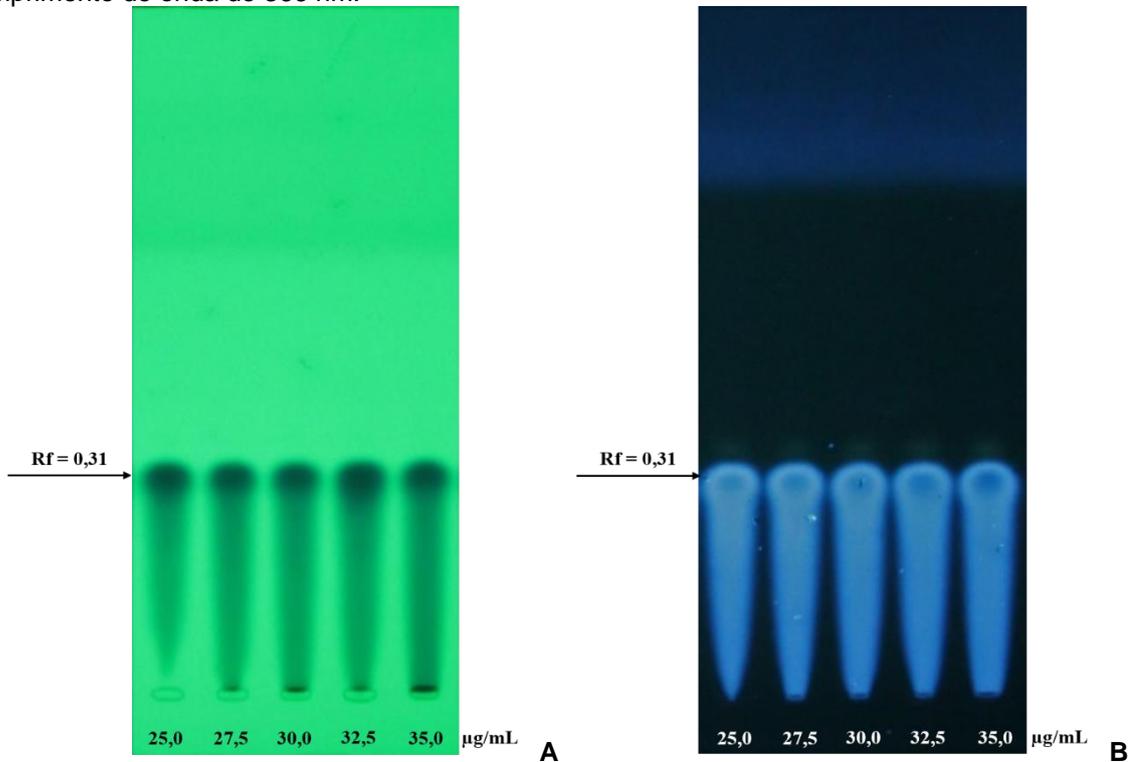


Fonte: Autoria própria.

4.1.4 Ácido cafeico

Para o padrão ácido cafeico foi encontrado Rf correspondente a 0,31. As bandas apresentaram-se arrastadas devido à alta concentração como pode ser visualizado na figura 10. Para tentar solucionar esse impasse e obter bandas mais definidas, menores concentrações foram testadas. Entretanto, foi verificado que a ferramenta de imagem não detectava as bandas de interesse, não sendo possível obter os dados de imagem. Dessa forma, a análise foi realizada empregando as concentrações na faixa de 25 – 35 $\mu\text{g/mL}$.

Figura 10 - Cromatoplasmas desenvolvidas para análise de imagem do padrão ácido cafeico. Faixa de concentração: 25-35 µg/mL. Sistema de eluição - Tolueno: AcOEt: HCOOH (4,5:3:0,2). A - Cromatoplasca visualizada sob o comprimento de onda de 254 nm; B – Cromatoplasca visualizada sob o comprimento de onda de 366 nm.

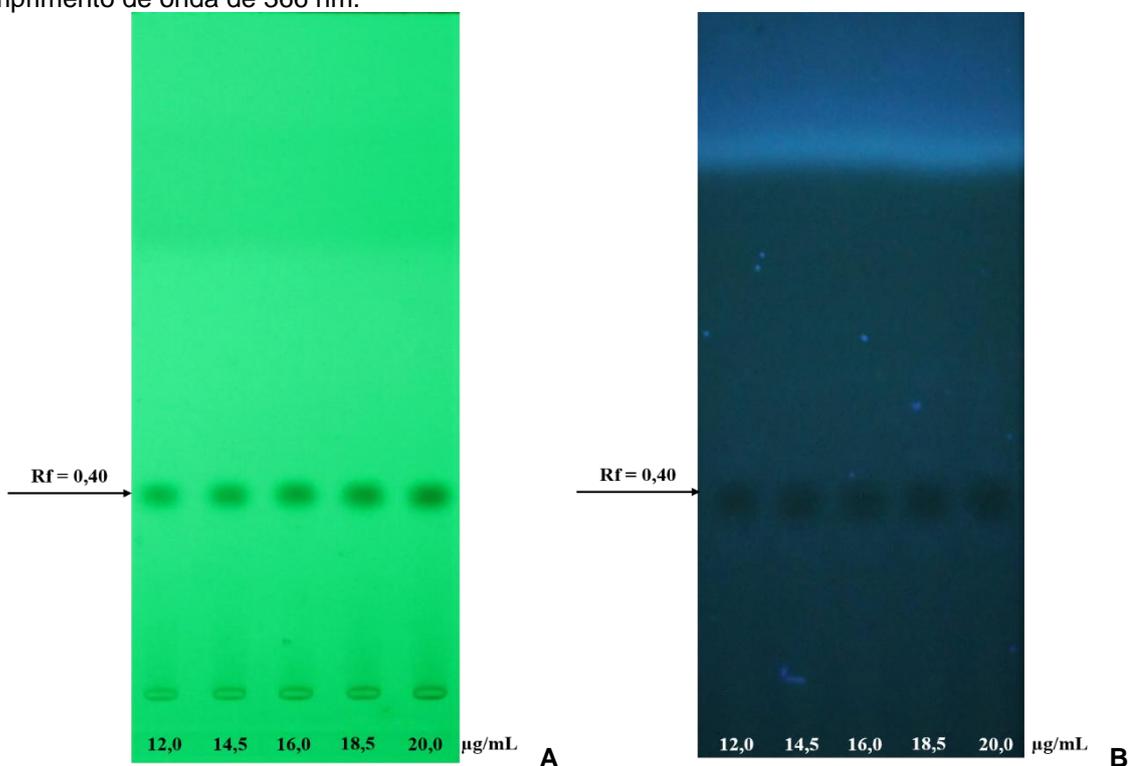


Fonte: Autoria própria.

4.1.5 Catequina

As imagens digitalizadas das cromatoplasmas do padrão catequina estão apresentadas na figura 11. O Rf das bandas do padrão correspondeu a 0,40, cujas formas apresentaram-se definidas e sem presença de arraste com o sistema de eluição empregado.

Figura 11 - Cromatoplasmas desenvolvidas para análise de imagem do padrão Catequina. Faixa de concentração: 12-20 $\mu\text{g/mL}$. Sistema de eluição - tolueno: AcOEt: HCOOH: MeOH (3:3:0,8:0,2). A - Cromatoplasma visualizada sob o comprimento de onda de 254 nm; B - Cromatoplasma visualizada sob o comprimento de onda de 366 nm.



Fonte: Autoria própria.

Os resultados anteriormente descritos tratam de adequações que foram moduladas utilizando a ferramenta de imagem ImageJ. Para análise no Sidtermo, as mesmas imagens foram utilizadas, sem necessidade de tratamento prévio.

4.1.6 *Schinus terebinthifolius*

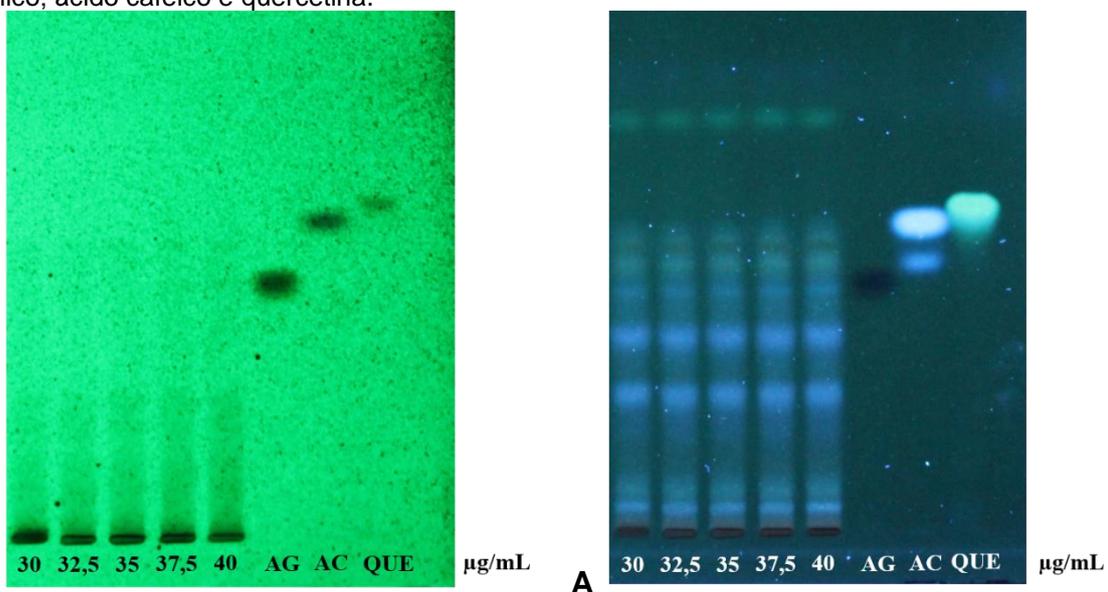
Para análise das cascas de *S. terebinthifolius* foram testados a matéria prima vegetal, o extrato bruto (EB) e a fração enriquecida (acetato de etila).

Inicialmente, foram testadas amostras das cascas de matéria prima vegetal. Foram aplicados 2 μL da solução obtida para a amostra na cromatoplasma, e, após eluição não foi detectada a presença do marcador de interesse na espécie, ácido gálico. Uma justificativa pela qual o composto fitoquímico não ter sido visualizado pode ter sido devido à concentração da solução empregada. A banda mostrou-se com arraste predominante de cor marrom na placa cromatográfica que não possibilitou detectar a presença do composto. Dessa forma, a alternativa foi utilizar o emprego de soluções mais diluídas. Da solução mãe (1 mg/mL) foi retirada uma alíquota de 100

μL e diluída em 1 mL de metanol e, posteriormente, a quantidade de 30 μL foi aplicada na cromatoplaça. Entretanto, não foi detectada a presença do tanino hidrolisável. Alíquotas maiores da solução mãe foram utilizadas, retirou-se a faixa corresponde a 200-500 μL e mesmo assim o marcador não foi detectado e à medida em que aumentava a concentração a banda apresentava arraste predominante.

Diante dos resultados observados, optou-se por trabalhar com a análise do marcador no extrato bruto. No entanto, o extrato bruto apresentou os mesmos impasses encontrados para matéria prima. Ao realizar análise da fração enriquecida acetato de etila foi possível detectar a presença do composto de interesse sob as condições estabelecidas na metodologia. Frações enriquecidas evidenciam maior extração de compostos fenólicos, devido à solubilidade em solventes de polaridade intermediária como o acetato de etila (BURQUE, 2015). A figura 12 mostra a cromatoplaça de aroeira. Foi realizada uma triagem fitoquímica os padrões de ácido gálico, ácido cafeico e quercetina com a finalidade de detectar esses polifenóis na espécie. O resultado foi positivo apenas para ácido gálico, cujo Rf foi igual a 0,42.

Figura 12 - Cromatoplaças desenvolvidas para análise de imagem da fração acetato de etila de *S. terebinthifolius*. Faixa de concentração: 30-40 $\mu\text{g/mL}$. Sistema de eluição - tolueno: AcOEt: HCOOH: MeOH (3:3:0,8:0,2). A – Cromatoplaça visualizada sob o comprimento de onda de 254 nm; B – Cromatoplaça visualizada sob o comprimento de onda de 366 nm. Padrões à direita da placa: ácido gálico; ácido cafeico e quercetina.



Onde: AG – Ácido gálico; AC – Ácido cafeico; QUE – Quercetina.
Fonte: Autoria própria.

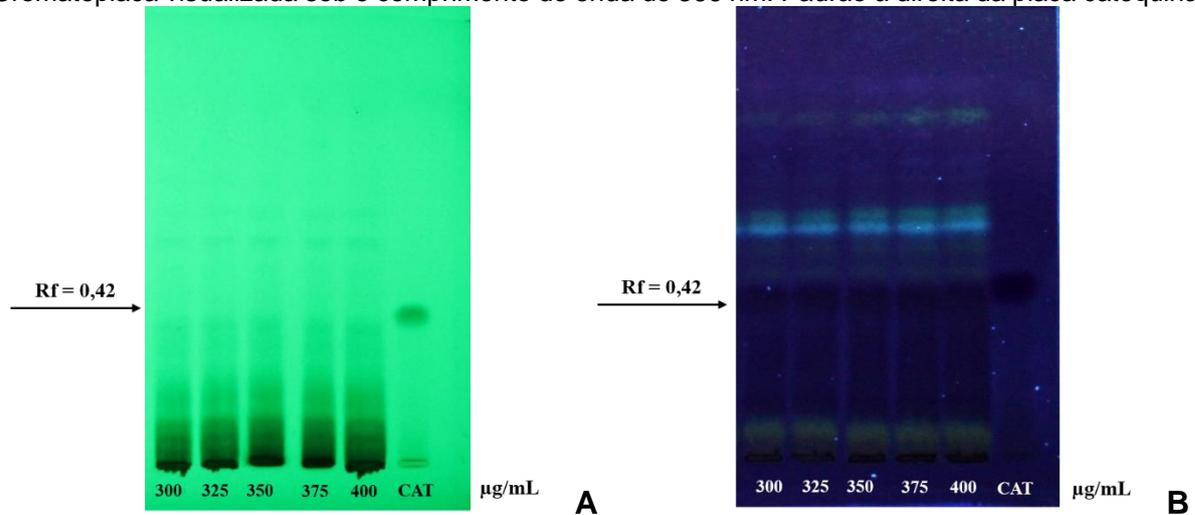
As imagens apresentaram condições satisfatórias para análise na ferramenta de imagem Sidtermo. O software ImageJ não apresentou sensibilidade para detectar a região de interesse.

4.1.7 *Anadenanthera colubrina*

Para o desenvolvimento das cromatoplasmas de *A. colubrina*, algumas limitações foram encontradas. Ao aplicar a faixa de concentração de 300 a 400 µg/mL da solução amostra, não foi possível detectar o marcador químico de interesse da espécie (catequina). Testes anteriores ao emprego dessa faixa de concentração foram realizados como, por exemplo, partindo-se de uma solução mãe mais concentrada. No entanto, foi encontrada dificuldade com a solubilização da amostra não sendo possível utilizar uma solução mais concentrada, devido a possíveis danos que poderiam causar à seringa de aplicação, além da dificuldade de aspiração. Ao aplicar soluções mais diluídas não foi possível verificar a presença de catequina no extrato bruto.

Assim, a alternativa foi a utilização de agente cromogênico (vanilina clorídrica) que garante a especificidade e propicia o aumento da intensidade da coloração. Ao derivatizar a cromatoplasma, com a faixa de concentração de 300 a 400 µg/mL, as bandas correspondentes ao marcador catequina foram visualizadas sob os dois comprimentos de onda (254 nm e 366 nm), cujo R_f correspondeu a 0,42. As placas cromatográficas reveladas e digitalizadas estão apresentadas na figura 13.

Figura 13 - Cromatoplasmas desenvolvidas para análise de imagem da fração acetato de etila de *A. colubrina*. Faixa de concentração: 300-400 µg/mL. Sistema de eluição - tolueno: AcOEt: HCOOH: MeOH (3:3:0,8:0,2). A – Cromatoplasca visualizada sob o comprimento de onda de 254 nm; B – Cromatoplasca visualizada sob o comprimento de onda de 366 nm. Padrão à direita da placa catequina.



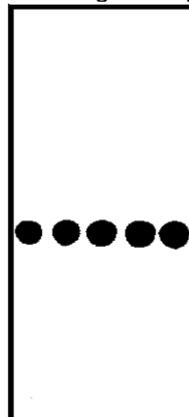
Fonte: Autoria própria.

As imagens apresentaram condições satisfatórias para análise na ferramenta de imagem Sidtermo. Assim como para a aroeira, o software ImageJ não apresentou sensibilidade para detectar de forma isolada a região de interesse, além disso, os resultados obtidos durante a configuração da rede neural artificial (próximas sessões) apontaram o Sidtermo como melhor ferramenta de análise.

4.2 CONFIGURAÇÃO DE REDE NEURAL ARTIFICIAL (RNA)

Após tratamento (Figura 14) e obtenção das medidas de imagem com as condições, foi realizada a correlação entre a área da banda e a concentração correspondente para verificar a proporcionalidade entre as variáveis.

Figura 14 - Imagem de cromatoplasca do padrão ácido gálico após processamento no software ImageJ.



Fonte: Autoria própria.

Inicialmente, foi realizada uma análise com a área da banda do padrão de ácido gálico. A tabela 1 mostra resultados de coeficiente de determinação (R^2) obtidos para o ácido gálico quando foi realizada a correlação da área da banda com a respectiva concentração. Verificou-se que a área da maioria das imagens não foi proporcional a concentração quando uma relação linear foi estabelecida, pois os coeficientes de determinação apresentaram valores inferiores a 0,99, cujo valor é indicativo de relação entre os dados (BRASIL, 2017).

Tabela 1 - Valores do coeficiente de determinação (R^2) para o padrão ácido gálico ao aplicar relação linear. Faixa de concentração: 20-30 $\mu\text{g/mL}$.

Placa cromatográfica	Valores de R^2
1	0,8347
2	0,8762
3	0,6651
4	0,9437
5	0,6080
6	0,7653
7	0,9056
8	0,7846
9	0,7423
10	0,9454

Fonte: Autoria própria.

O emprego das RNAs funciona como uma ferramenta essencial para resolução de problemas que requerem análise de modelos não lineares, sendo capaz de estabelecer correlação entre variáveis que não apresentam proporcionalidade (KUSTRIN, 2013).

Devido à baixa associação entre a área da banda com a respectiva concentração verificada através do coeficiente de determinação, a alternativa para relacionar os dados das imagens à concentração foi empregar RNA, por ser uma técnica que apresenta como característica a robustez e tolerância à falha, produzindo resultados confiáveis, além de não se restringir a modelos lineares. As pesquisas em inteligência computacional reorganizam o processamento de dados considerando informações paralelas, não ficando limitado a modelos lineares e sequenciais (ALMEIDA et al., 2009).

Nesse sentido, o modelo Kernel é empregado para mapear os dados de entrada no espaço de recursos de alta dimensionalidade. Assim, os diferentes tipos de modelo Kernel foram avaliados com a finalidade de explorar as relações complexas entre os dados e dessa forma estimar a configuração ideal para determinar a concentração de polifenóis a partir de medidas de imagem.

O método Kernel tenta modelar a relação de variável resposta denominada de y (concentração) e variável explicativa conhecida como x (dados de imagem). Nem sempre será estabelecida uma relação linear entre as variáveis explicativas e variáveis de resposta sendo necessário avaliar métodos alternativos a regressão linear (DOUGLAS, 2012).

Para estimar a configuração de rede foram avaliados Kernel linear, sigmoide, polinomial, gaussiano (RBF), dilatação e erosão. Todos esses modelos têm a abordagem clássica de treinar uma rede neural artificial e calcular a matriz de pesos (MOTA et al., 2011). Com exceção do Kernel linear, todos os outros modelos empregados fundamentam-se na teoria não linear na perspectiva de modelar os dados de imagem com a respectiva concentração. Os modelos de Kernel foram avaliados de acordo com os valores de erro percentual médio (EPM), cuja variável indica a eficiência de tratamento dos dados, quanto menor o erro percentual médio melhor desempenho em treinar uma rede, uma vez que os resultados se aproximam do valor ideal. Dessa forma, serão apresentados e analisados os resultados obtidos para cada modelo testado com base no EPM dos padrões ácido gálico, quercetina, rutina, ácido cafeico e catequina.

4.2.1 Ácido gálico

4.2.1.1 Análise de imagem por ImageJ

Para o padrão ácido gálico em 254 nm, no teste divisão percentual foi verificado que os modelos sigmoide, linear e polinomial (P2) apresentaram os menores valores de erro percentual médio correspondendo respectivamente a 0,018; 0,007; e, 0,040, indicando que esses modelos representam a configuração mais adequada no processo de modelagem para correlacionar as medidas de imagem com a concentração do padrão de ácido gálico. Não existe limite de especificação para determinar o EPM, no entanto, quando modelos são comparados entre si, o que indica

menor valor apresenta melhor desempenho para modelagem de dados. Os modelos polinomiais de grau 3, 4 e 5 apresentaram valores acima de 0,7 como pode ser verificado na tabela 2. Portanto, os modelos inicialmente citados representam configurações com melhores ajustes para correlação de dados provenientes de medidas de imagem com a quantificação do padrão. Ao realizar a análise do modelo gaussiano (RBF) foi obtido valor de EPM superior a 1, indicando que essa configuração não é ideal para estabelecer uma associação entre dados de imagens com a concentração do analito devido ao alto erro apresentado.

Tabela 2 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 254 nm com atributos de imagem gerados pelo ImageJ.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,018	0,007	0,040	0,709	0,713	0,719	1,000	71,290	14,466
Validação cruzada	0,025	0,009	0,038	0,764	0,697	0,713	1,000	70,989	35,497

Onde: S -Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.
Fonte: Autoria própria.

Uma Rede Neural de Base Radial possui nós de entrada e apenas duas camadas, uma corresponde à camada intermediária e outra a camada de saída, cujo processo de aprendizagem é baseado na teoria da programação não linear (MOTA et al., 2011).

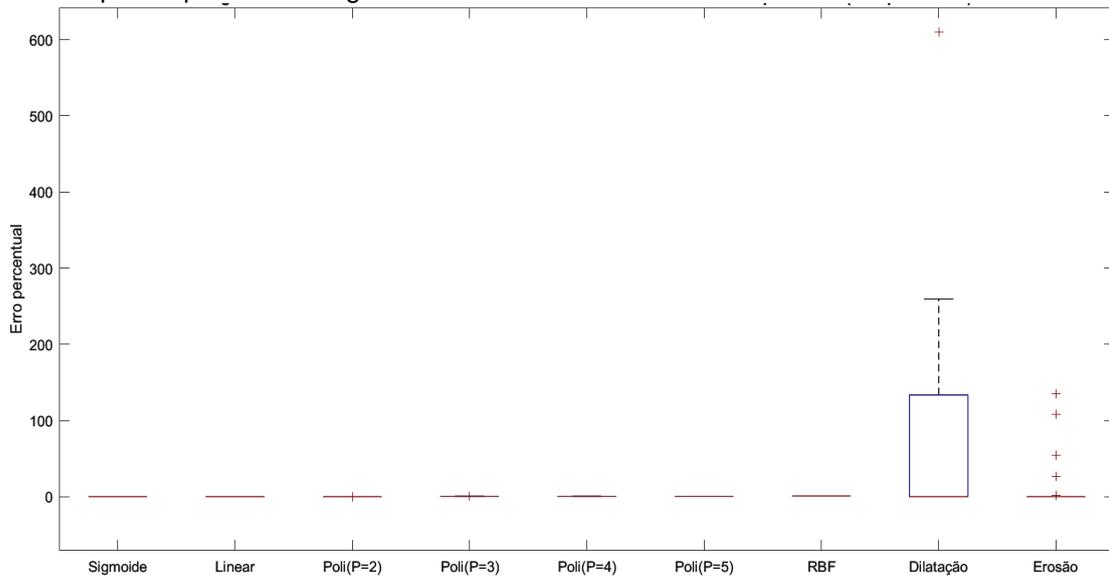
Para os modelos dilatação e erosão foram verificados os maiores valores de erro percentual médio, logo esses modelos não são considerados ideais para a correlação de medidas de imagem com a concentração do padrão em análise.

Com relação ao teste validação cruzada realizado para as imagens em 254 nm foi verificado maior desempenho para os modelos sigmoide, linear e polinomial 2 devido ao menor valor de erro percentual médio correspondendo respectivamente a 0,025; 0,009 e 0,038. Portanto, esses modelos foram considerados mais apropriados para correlacionar atributos de imagem do padrão ácido gálico com a respectiva

concentração testada. Além disso, foi verificado que os modelos dilatação e erosão apresentaram EPM muito elevado não sendo ideais para estabelecer configuração de rede.

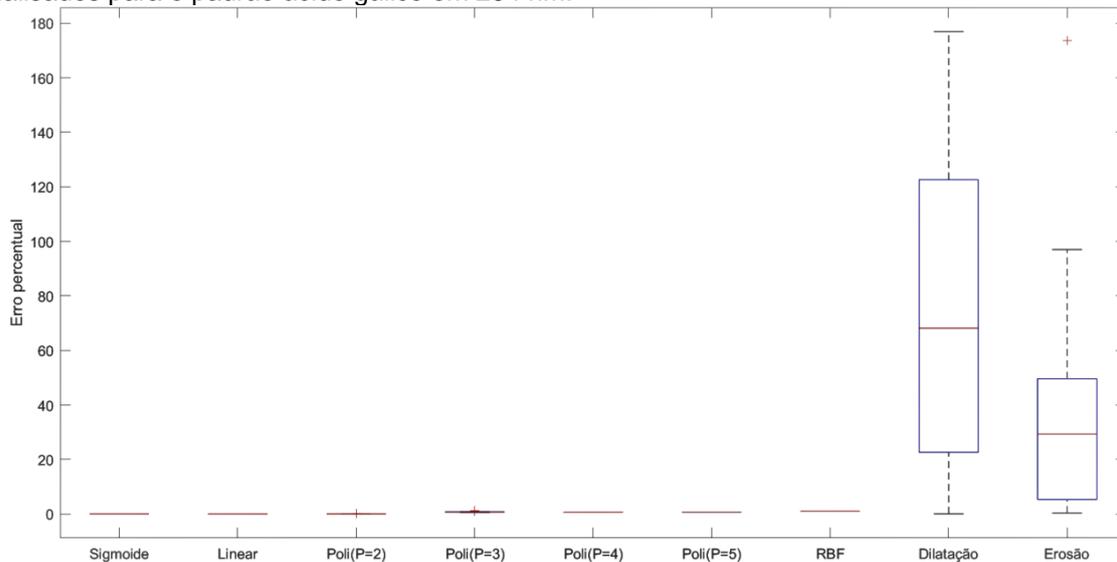
Para melhor visualização e interpretação estatística dos resultados, os dados foram expressos pelo gráfico do tipo Boxplot (Figuras 15 e 16), onde o EPM e o desvio padrão de cada modelo são evidenciados. O gráfico permite visualizar o alto erro percentual e desvio padrão entre as curvas de calibração analisadas para os modelos dilatação e erosão quando o teste de divisão percentual e validação cruzada foram realizados. Esse fato é justificado pela rigidez do modelo a ponto de não conseguir modelar fielmente os dados (BINOTI et al., 2014). Os sinais *outliers* (+) mostrados no gráfico representam os valores desgarrados, indicando que algumas curvas de calibração apresentaram valores fora do padrão estabelecido pelo sistema quando comparadas entre si. Essa variação atípica pode ter sido decorrente de erro de leitura, erro na execução do experimento ou característica inerente à própria variável (CALLEGARI-JACQUES, 2003).

Figura 15 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 254 nm.



Fonte: Autoria própria.

Figura 16 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 254 nm.



Fonte: Autoria própria.

Ao realizar o teste de divisão percentual com as imagens digitalizadas em 366 nm, foi verificado que os modelos dilatação e erosão apresentaram maior desempenho na correlação dos atributos de imagem com a quantificação do analito. O mesmo foi observado quando o teste de validação cruzada foi realizado. Os valores estão apresentados na tabela 3.

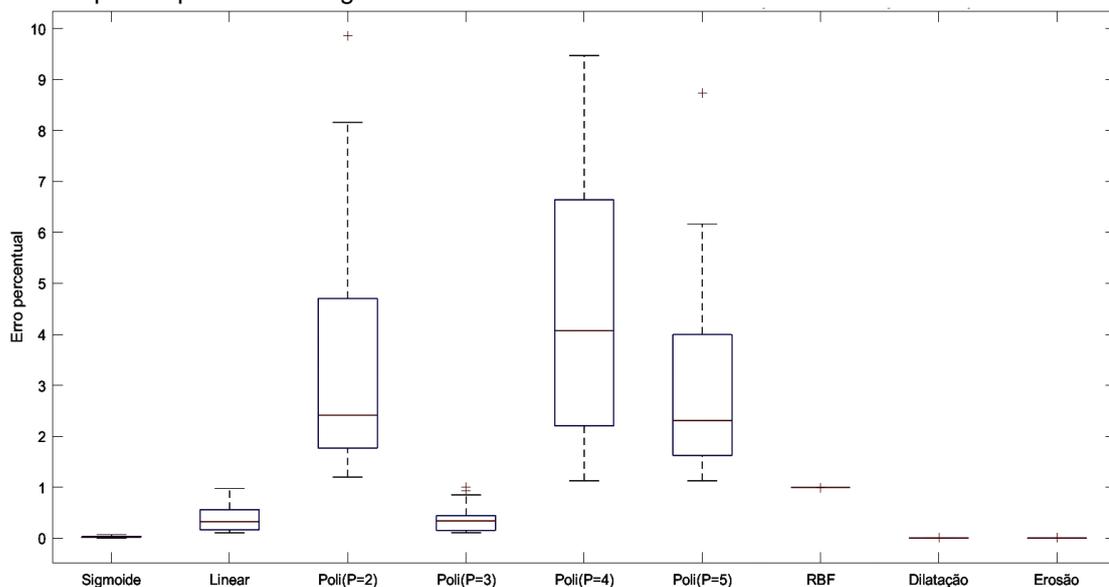
Tabela 3 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 366 nm com atributos de imagem gerados pelo ImageJ.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,027	0,015	1,152	0,930	1,930	3,560	0,998	0,002	0,002
Validação cruzada	0,027	0,008	2,198	2,789	6,172	5,261	0,996	0,003	0,003

Onde: S-Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.
 Fonte: Autoria própria.

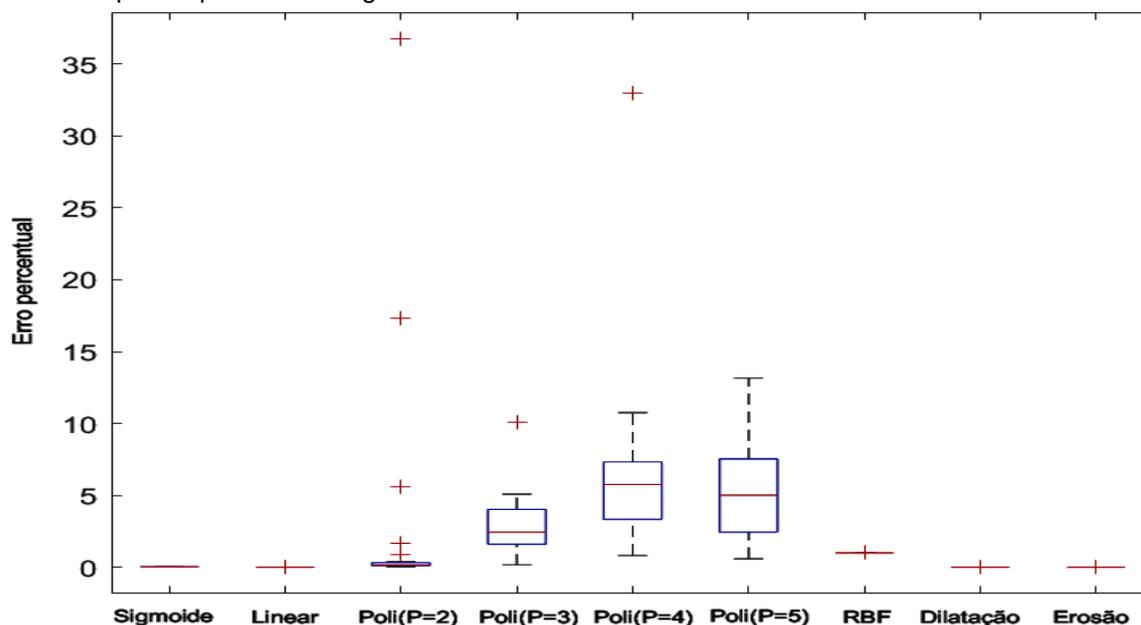
O gráfico de Boxplot (Figura 17) para o teste divisão percentual das imagens tratadas sob o comprimento de onda 366 nm evidencia alto desvio padrão para os modelos polinomial 2, 4 e 5. Além disso, é possível visualizar que os modelos sigmoide, dilatação e erosão apresentaram menor desvio padrão e menor erro percentual médio. Para o teste validação cruzada foram verificadas curvas de calibração fora do padrão estimado quando o teste foi realizado (*outliers*) ao empregar os modelos polinomial 2, 3 e 4 (Figura 18). Apresentaram bom desempenho para correlação dos dados os modelos sigmoide, linear, dilatação e erosão.

Figura 17 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 366 nm.



Fonte: Autoria própria.

Figura 18 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 366 nm.



Fonte: Autoria própria.

4.2.1.2 Análise de imagem por Sidtermo

Com a finalidade de encontrar a ferramenta de imagem que possibilitasse melhor aplicabilidade de desempenho para análise das bandas cromatográficas outro software de imagem foi testado. O *software* Sidtermo permitiu análises mais rápidas que o ImageJ e não foi necessário selecionar a região de interesse, uma vez que o programa faz a análise levando em consideração toda imagem. Cada concentração aplicada na cromatoplaça foi analisada individualmente para evitar interferência das demais concentrações nas análises. Dessa forma, os atributos de imagem gerados pela ferramenta são precisos para cada concentração empregada.

A tabela 4 apresenta os valores de erro percentual médio para os testes divisão percentual e validação cruzada com os atributos gerados pelo Sidtermo ao analisar as cromatoplaças sob o comprimento de onda de 254 nm. Realizando a comparação das duas ferramentas de imagem ao aplicar o teste divisão percentual é possível verificar que para maioria dos modelos testados, os menores EPM foram observados quando a análise das imagens foi realizada pelo Sidtermo, excetuando-se os modelos de Kernel sigmoide, linear e polinomial 2. Dentre os modelos com menor EPM considera-se que melhor desempenho foi obtido com o modelo sigmoide. Para o teste de validação cruzada, ao realizar a comparação dos dados provenientes das duas

ferramentas de imagem, foi possível constatar os mesmos níveis de desempenho para os modelos. O Kernel polinomial 3 apresentou o melhor desempenho para correlação de atributos de imagem com a respectiva quantificação do analito, cujo erro percentual médio foi igual a 0,194.

Tabela 4 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 254 nm com atributos de imagem gerados pelo Sidtermo.

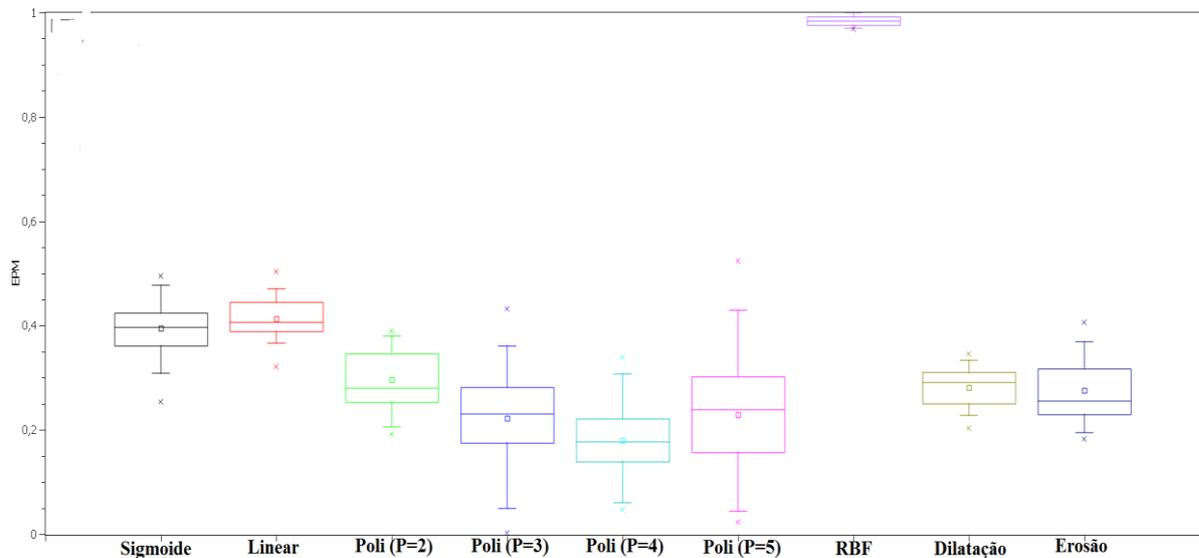
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,039	0,413	0,295	0,222	0,179	0,229	0,983	0,280	0,275
Validação cruzada	0,395	0,407	0,273	0,194	0,208	0,221	0,984	0,284	0,286

Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

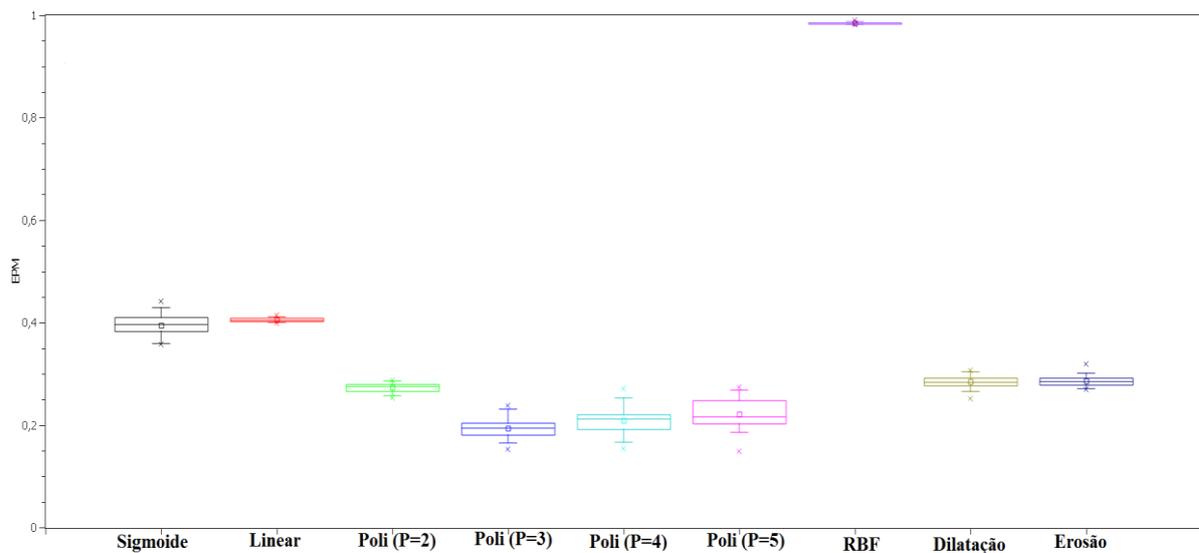
Os gráficos de Boxplot para ambos os testes aplicados para as cromatoplasmas registradas sob o comprimento de 254 nm estão apresentados nas figuras 19 e 20. Para o teste divisão percentual, o gráfico mostra que o desvio padrão para todos os modelos analisados foram maiores quando comparados aos dados gerados pelos atributos de imagem do ImageJ. Possivelmente, isso é justificado porque não há interferência de tratamento das imagens que possa minimizar ruídos que estejam presentes na imagem. Para o teste validação cruzada, verifica-se menor desvio padrão para todas as variáveis de Kernel. Assim, esse segundo teste apresenta resultados mais precisos quando comparado ao teste de divisão percentual.

Figura 19 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 254 nm.



Fonte: Autoria própria.

Figura 20 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 254 nm.



Fonte: Autoria própria.

Para o padrão ácido gálico sob o comprimento de onda de 254 nm, ficam estabelecidas como as melhores condições, para correlação dos dados de imagem gerados pelo Sidtermo com a respectiva concentração, o teste validação cruzada com os modelos polinomiais de grau 2 e 3.

Ao fazer análise comparativa entre as ferramentas de imagem (ImageJ e Sidtermo) para os dados provenientes das cromatoplasmas digitalizadas sob o comprimento de 366 nm é possível verificar que os modelos kernel polinomial 3, 4 e 5 apresentaram melhores desempenhos para o teste divisão percentual quando os

atributos de imagem foram gerados pelo Sidtermo. Isso também foi evidenciado para o teste de validação cruzada (Tabela 5).

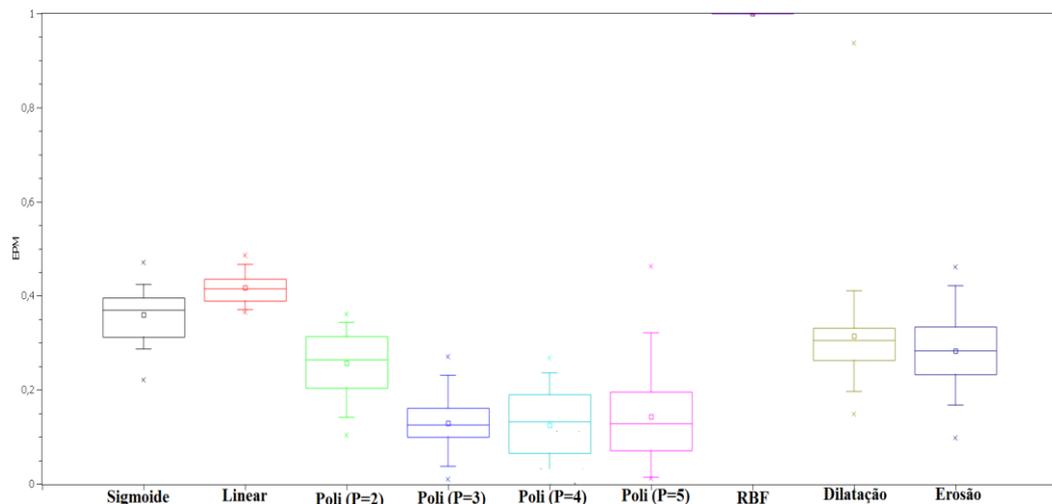
Os Boxplots de ambos os testes empregados para as imagens registradas sob o comprimento de onda de 366 nm estão apresentados nas figuras 21 e 22. De acordo com o gráfico para o teste divisão percentual é possível verificar que a maioria dos modelos apresentou alto desvio padrão. Enquanto, para o teste validação cruzada, o contrário é verificado, o desvio padrão apresentou-se relativamente menor para os modelos analisados, representando maior acurácia quando esse tipo de teste é aplicado para modelagem dos dados.

Tabela 5 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido gálico em 366 nm com atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,035	0,417	0,256	0,129	0,124	0,143	1,000	0,314	0,285
Validação cruzada	0,378	0,411	0,255	0,163	0,180	0,181	1,000	0,300	0,298

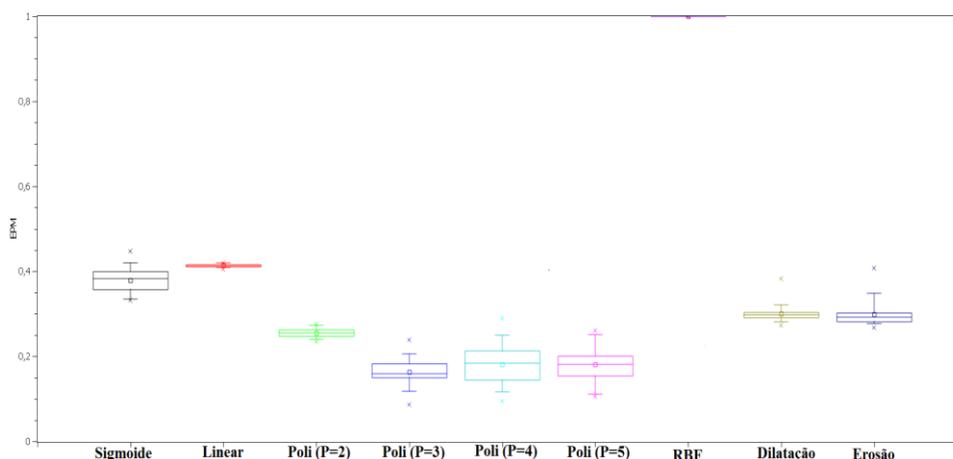
Onde: S-Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.
Fonte: Autoria própria.

Figura 21 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido gálico em 366 nm.



Fonte: Autoria própria.

Figura 22 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido gálico em 366 nm.



Fonte: Autoria própria.

4.2.2 Quercetina

4.2.2.1 Análise de imagem por ImageJ

Ao realizar a análise com o padrão em 254 nm para o teste divisão percentual foi verificado que os modelos linear, dilatação e erosão apresentaram resultados mais adequados para a finalidade proposta, cujos valores de EPM correspondem respectivamente a 0,0033; 0,0028 e 0,0025. Estes achados indicam que esses modelos são ideais para correlação dos dados de imagem com a concentração do padrão. Os valores de EPM para os demais modelos em estudo estão evidenciados na tabela 6.

Tabela 6 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 254 nm com atributos de imagem gerados pelo ImageJ.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,072	0,003	0,019	0,069	0,455	0,293	1,000	0,002	0,002
Validação cruzada	0,072	0,003	0,049	0,160	0,513	0,206	1,000	0,003	0,003

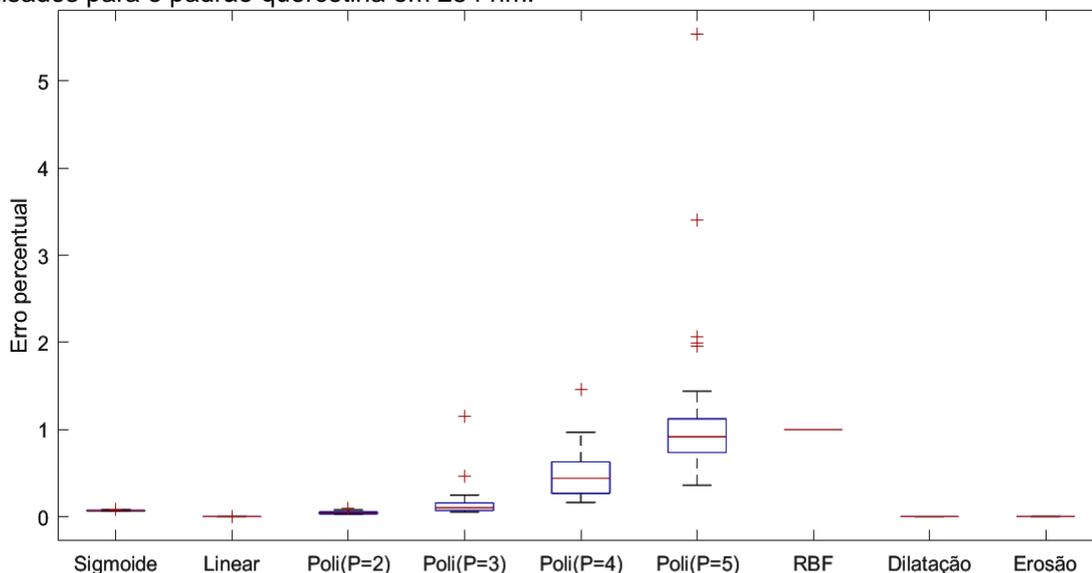
Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

Para o teste validação cruzada para as cromatoplasmas em 254 nm foi verificado menor erro para o Kernel linear, dilatação e erosão sendo encontrado o mesmo valor de EPM para os três modelos estudados. O Kernel sigmoide e polinomial 2 apresentou EPM mais elevado quando comparado aos primeiros citados, no entanto também indicou bom desempenho para correlação dos dados de imagem com a concentração de quercetina.

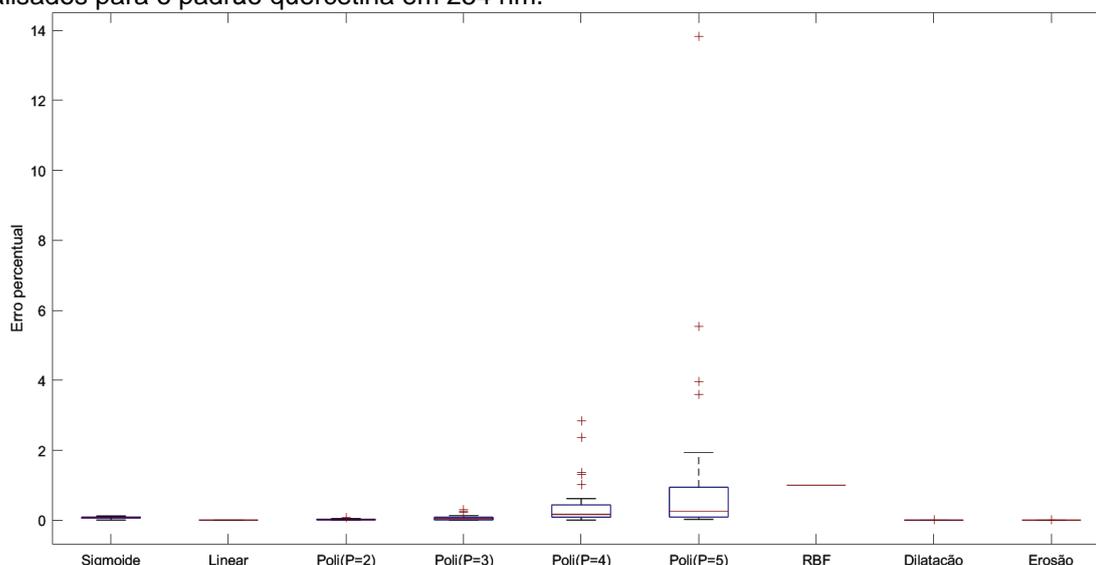
O Boxplot para ambos os testes (divisão percentual e validação cruzada) pode ser visualizado nas figuras 23 e 24. Para divisão percentual é possível visualizar que o modelo polinomial 4, 5 e RBF apresentaram os maiores valores de EPM, logo essas configurações não são adequadas para correlacionar os valores de imagem com a quantificação do fitoconstituente. Além disso, foi verificado os maiores valores de desvio padrão e sinais de *outliers* (+) para o Kernel polinomial 3, 4 e 5. O modelo polinomial 5 foi o que evidenciou maior discrepância entre os dados das curvas de calibração não correspondendo a um modelo ideal para modelagem de rede. Para o teste validação cruzada foi verificado maior desvio padrão e sinais *outliers* (+) para Kernel polinomial 4 e 5.

Figura 23 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão quercetina em 254 nm.



Fonte: Autoria própria.

Figura 24 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão quercetina em 254 nm.



Fonte: Autoria própria.

Em 366 nm para o padrão quercetina tanto para o teste de divisão percentual quanto para validação cruzada foi verificado que o modelo linear, dilatação e erosão apresentaram maior desempenho, cujos valores de erro percentual médio estão apresentados na tabela 7.

Tabela 7 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 366 nm com atributos de imagem gerados pelo ImageJ.

Teste	Modelos								
	S	Linea r	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentua l	0,091	0,022	0,367 0	0,857 8	2,23 8	4,45 8	1,00 0	0,02 9	0,01 9
Validação cruzada	0,120 5	0,028	0,215	1,417	1,25 2	2,49 2	1,00 0	0,02 8	0,03 1

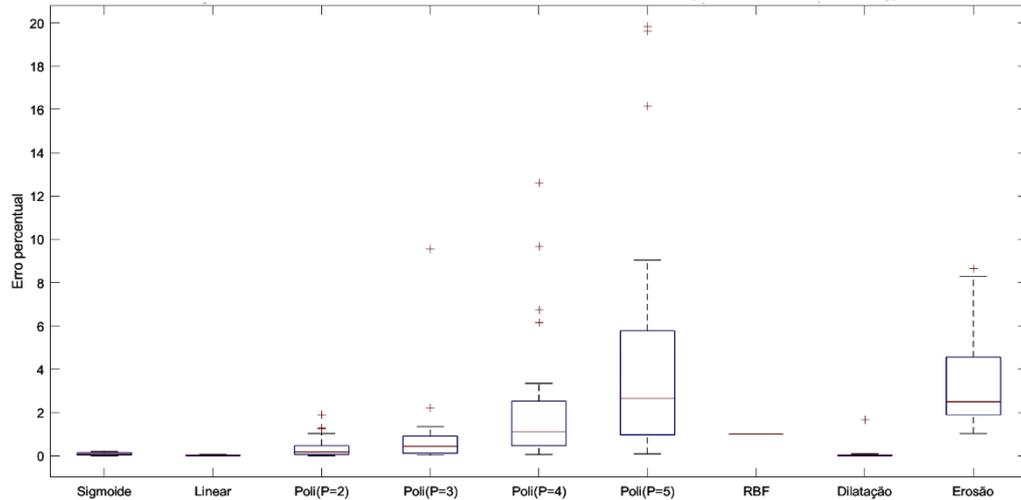
Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

Apesar do modelo erosão para o teste divisão percentual ter apresentado baixo EPM foi verificado alto desvio padrão quando o gráfico de Boxplot (Figura 25) foi

analisado. Além disso, foi verificado também que os modelos polinomiais apresentaram alto desvio padrão e muitos sinais *outliers* indicando que não são bons modelos para modelagem dos dados.

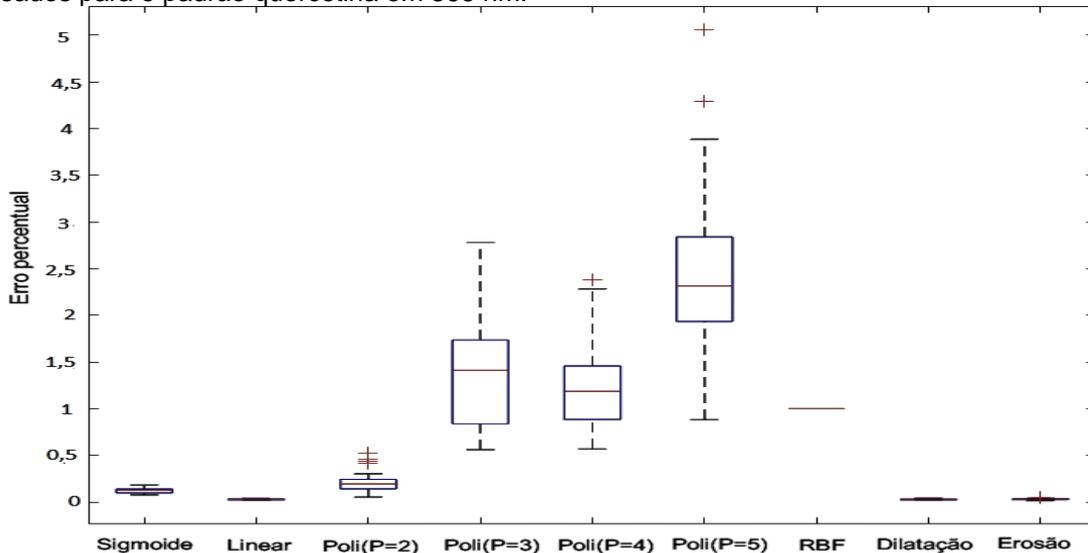
Figura 25 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão quercetina em 366 nm.



Fonte: Autoria própria.

A figura 26 evidencia de forma mais clara os resultados de erro percentual médio apresentados na tabela 5, além de apontar menor desvio padrão para os modelos com melhor desempenho para correlação dos dados de imagem do padrão quercetina com sua respectiva concentração.

Figura 26 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão quercetina em 366 nm.



Fonte: Autoria própria.

4.2.2.2 Análise de imagem Sidtermo

A tabela 8 apresenta os resultados dos EPMs obtidos para o padrão quercetina registrado sob o comprimento de 254 nm para os testes divisão percentual e validação cruzada dos atributos de imagem gerados pelo Sidtermo. Diante dos resultados verifica-se que para o teste divisão percentual os modelos polinomiais para os graus 3, 4 e 5 foram os que indicaram melhor desempenho. O modelo de base radial foi o que estipulou maior valor de EPM, cujo dado corresponde 0,999. O mesmo comportamento de resultados foi verificado quando o teste de validação cruzada foi aplicado, com valores relativamente próximos ao teste de divisão percentual.

Tabela 8 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 254 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,272	0,314	0,226	0,142	0,106	0,130	0,999	0,204	0,196
Validação cruzada	0,281	0,310	0,221	0,139	0,152	0,158	0,999	0,205	0,201

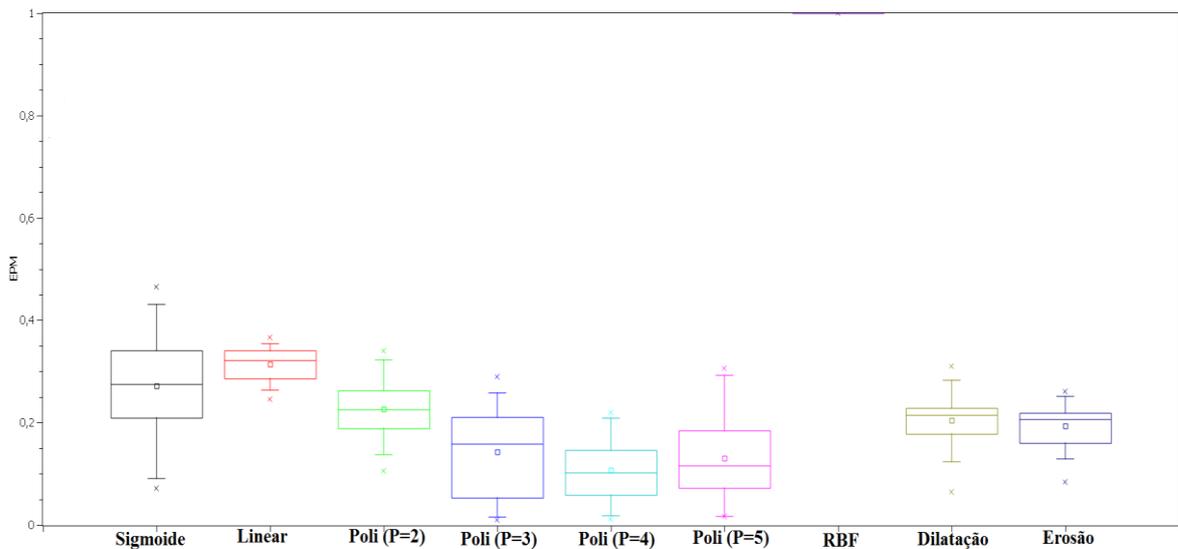
Fonte: Autoria própria.

Os gráficos de Boxplots para ambos os testes com as cromatoplasas digitalizadas sob o comprimento de onda de 254 nm estão apresentados nas figuras 27 e 28. Através da análise do gráfico (divisão percentual) é possível verificar que os modelos com maior desvio padrão são: sigmoide e os polinômios de grau 3 e 5. Os demais modelos apresentaram valores de desvio padrão menores. Dessa forma, ao analisar simultaneamente os valores encontrados para o EPM e o gráfico de Boxplot, o modelo polinomial 4 ($P = 4$) para o teste divisão percentual é escolhido como o de melhor desempenho para correlação dos atributos de imagem gerados pelo SidTermo com a respectiva faixa de concentração empregada.

Para o teste de validação cruzada o gráfico de bloxplot revela que o polinômio de grau 3 apresentou considerável desvio padrão quando comparado aos demais modelo de Kernel. Assim, o modelo polinomial 3 ($P = 3$) não é escolhido como ideal

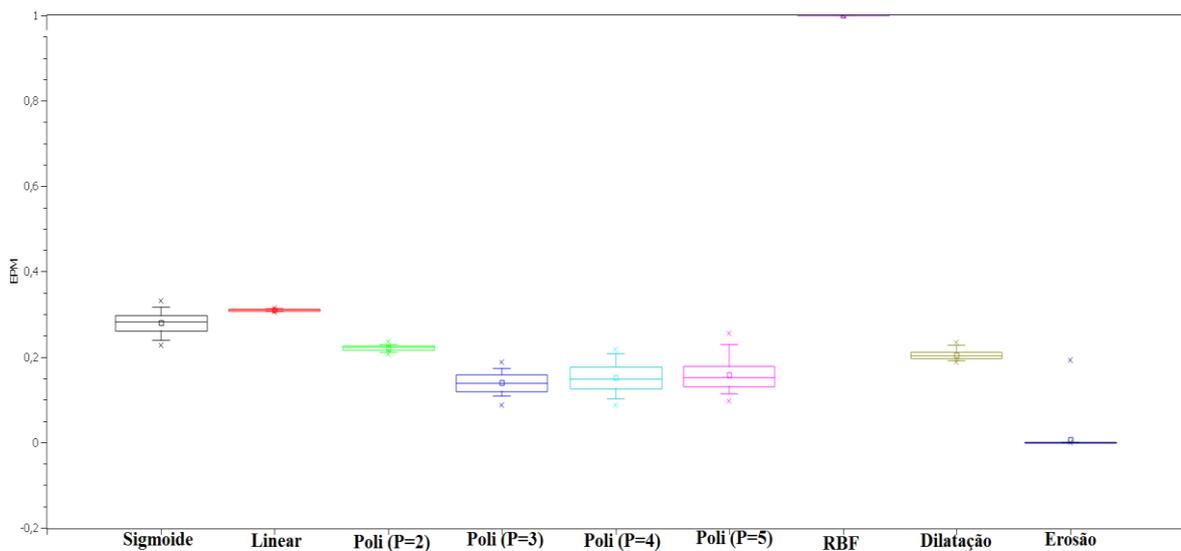
para modelagem dos dados mesmo apresentando baixo valor de EPM (0,194), pois o elevado desvio padrão apresentado aponta para falta de precisão de resultado. Dessa forma, os modelos de polinômios de grau 4 e 5 indicam melhor desempenho na modelagem de dados devido aos menores valores de EPMs e baixo desvio padrão para o teste de validação cruzada.

Figura 27 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão quercetina em 254 nm.



Fonte: Autoria própria.

Figura 28 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão quercetina em 254 nm.



Fonte: Autoria própria.

Os resultados para o padrão quercetina com as cromatoplasmas registradas sob o comprimento de onda de 366 nm estão apresentados na tabela 9. Para o teste de

divisão percentual, os modelos polinomiais 3, 4 e 5 apresentaram os menores erros percentuais médios, cujos valores correspondem, respectivamente, a 0,129; 0,112 e 0,105. Para o teste de validação cruzada, esses modelos também apresentaram os menores erros.

Tabela 9 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão quercetina em 366 nm dos atributos de imagem gerados pelo Sidtermo.

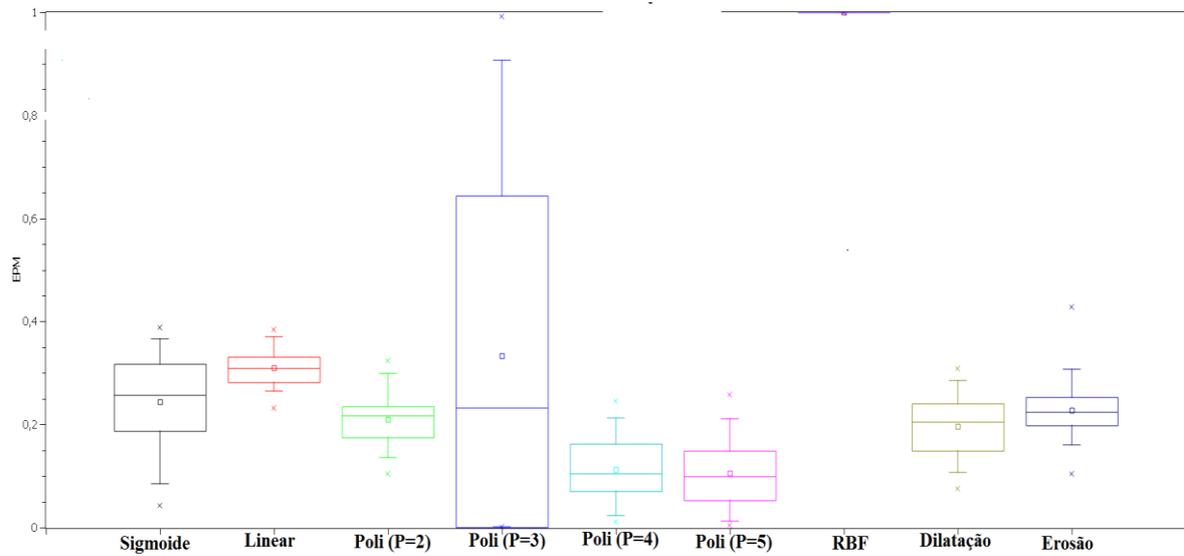
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,244	0,309	0,210	0,129	0,112	0,105	1,000	0,196	0,227
Validação cruzada	0,217	0,303	0,192	0,106	0,111	0,117	1,000	0,202	0,201

Fonte: Autoria própria.

Para auxílio na escolha do modelo Kernel que indique maior desempenho, os gráficos de Boxplots foram analisados (Figuras 29 e 30). O modelo polinomial 3 apresentou desvio padrão elevado quando o teste de divisão percentual foi aplicado. Os demais modelos também apresentaram alto desvio padrão.

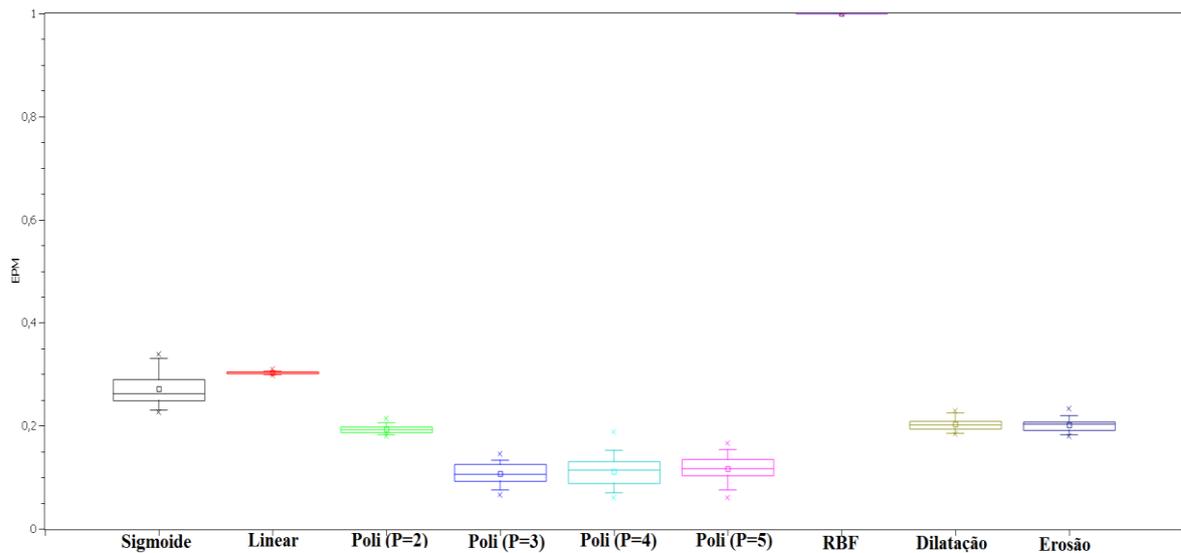
Para o teste validação cruzada, é possível verificar baixo desvio padrão para todos os modelos analisados. Praticamente não foi observado desvio padrão para o modelo de Kernel linear, entretanto devido ao maior EPM quando comparado aos demais modelos, esse modelo não foi escolhido como ideal para correlacionar os atributos de imagem com a concentração em estudo. Os polinômios de grau 3,4 e 5 apresentam-se como melhores para correlação de dados. Assim, o teste de validação cruzada com os polinômios com os melhores desempenhos e a ferramenta de imagem Sidtermo são as condições escolhidas para o desenvolvimento de rede neural artificial.

Figura 29 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão quercetina em 366 nm.



Fonte: Autoria própria.

Figura 30 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão quercetina em 366 nm.



Fonte: Autoria própria.

4.2.3 Rutina

4.2.3.1 Análise de imagem por ImageJ

Para as imagens no comprimento de onda em 254 nm ao empregar o teste divisão percentual foi verificado que os modelos linear, dilatação e erosão apresentaram os menores valores de erro percentual médio (Tabela 10). Estes valores de EPM indicam que tais modelos representam a configuração com maior desempenho para associar as medidas de imagem com as respectivas concentrações

do analito empregadas para preparar as cromatoplasacas. Para o teste validação cruzada foi verificado que o Kernel linear, dilatação e erosão também apresentaram os menores e semelhantes valores de EPM.

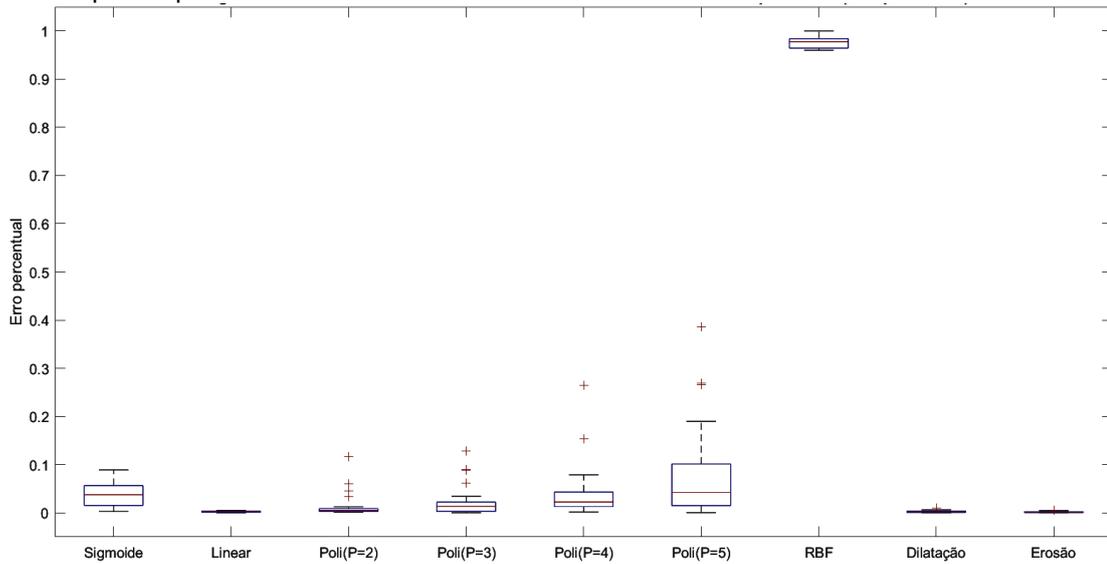
Tabela 10 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 254 nm com atributos de imagem gerados pelo ImageJ.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,038	0,002	0,013	0,022	0,037	0,079	0,977	0,002	0,002
Validação cruzada	0,073	0,003	0,009	0,021	0,064	0,122	0,974	0,003	0,003

Onde: S-Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.
Fonte: Autoria própria.

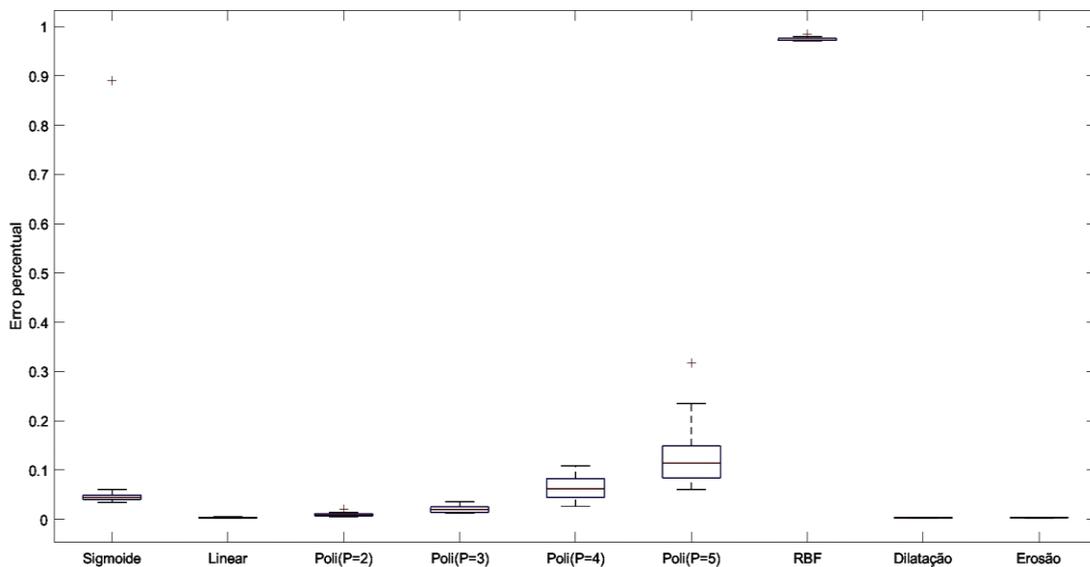
Para melhor interpretação dos resultados as figuras 31 e 32 mostram os gráficos de Boxplot para cada teste empregado possibilitando realizar análise estatística dos dados. Para o teste divisão percentual os modelos de Kernel sigmoide, RBF e polinomial 3, 4 e 5 apresentaram os maiores valores EPM quando foi realizada a associação dos atributos de imagem provenientes do padrão de rotina com suas respectivas concentrações. Estes resultados indicam que estes modelos não são os mais adequados para estimar configuração de rede quando comparado aos demais modelos que apresentaram menor valor de EPM. Os sinais *outliers* (+) mostrados no gráfico representam os pontos fora da curva, indicando que algumas curvas de calibração apresentaram valores fora do padrão quando comparadas uma com a outra.

Figura 31 - Boxplot das imagens tratadas pelo ImageJ para o teste de divisão percentual dos modelos analisados para o padrão rotina em 254 nm.



Fonte: Autoria própria.

Figura 32 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão rotina em 254 nm.



Fonte: Autoria própria.

A tabela 11 apresenta os resultados para as imagens visualizadas sob o comprimento de onda de 366 nm. De acordo com os valores de EPM obtidos é possível afirmar que os modelos que melhor correlacionam os dados de imagem provenientes do software ImageJ com a quantidade de rotina são o linear, polinomial 2, dilatação e erosão, cujos valores de EPM são os menores obtidos dentre os modelos analisados.

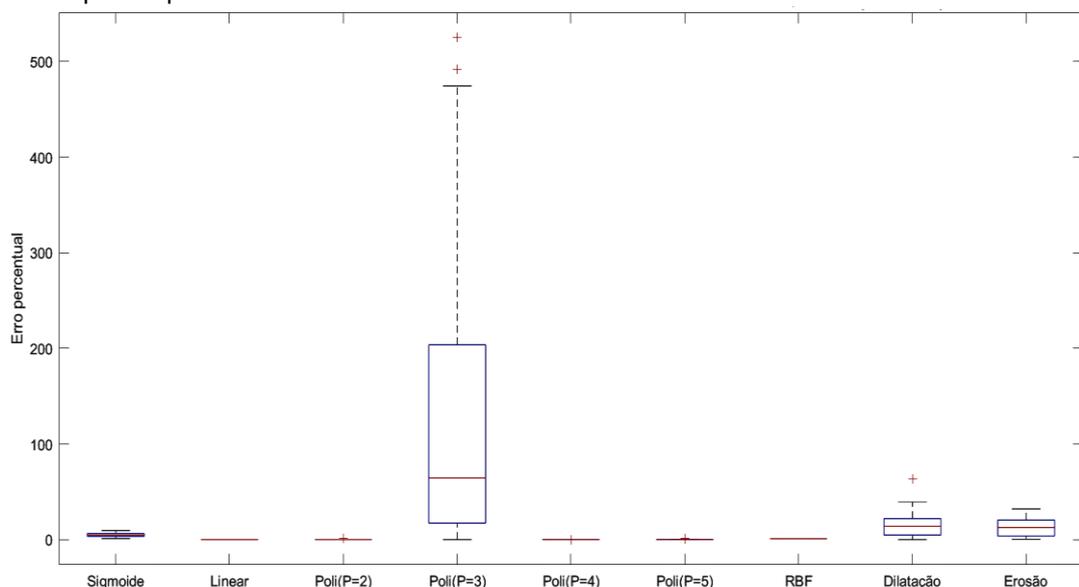
Tabela 11 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 366 nm com atributos de imagem gerados pelo ImageJ.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	1,543	0,001	0,008	0,016	0,041	0,133	1,000	0,002	0,001
Validação cruzada	0,044	0,002	0,006	0,012	0,038	0,096	1,000	0,002	0,002

Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.
Fonte: Autoria própria.

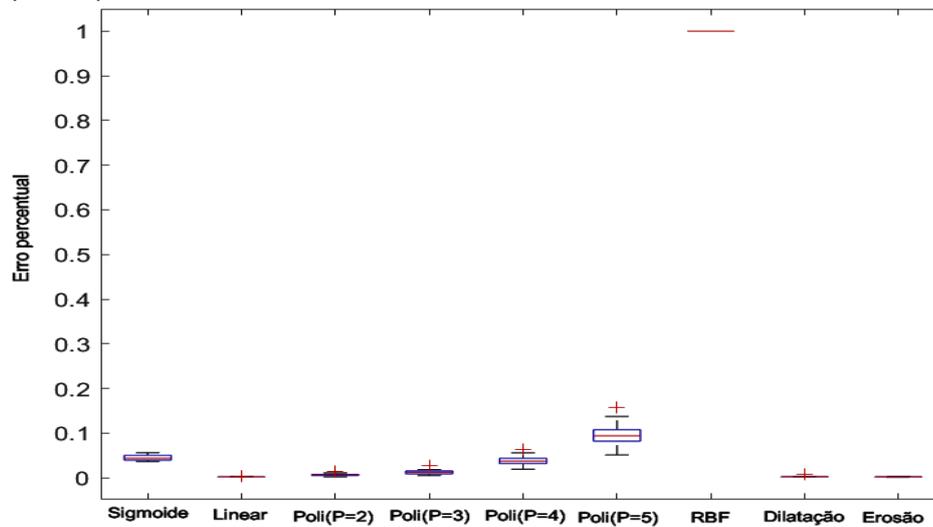
A figura 33 mostra o gráfico para o teste de divisão percentual empregado para análise do padrão de rotina sob o comprimento de 366 nm. De todos os modelos o polinomial apresentou-se mais desgarrado frente aos demais com relação ao desvio padrão e presença de curvas fora do padrão quando o teste foi aplicado. Ao realizar a análise do gráfico para o teste validação cruzada (Figura 34) em 366 nm foi verificado baixo desvio padrão para os modelos analisados e ausência de sinais *outliers* (+).

Figura 33 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão rotina em 366 nm.



Fonte: Autoria própria.

Figura 34 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão rutina em 366 nm.



Fonte: Autoria própria.

4.2.3.2 Análise de imagem Sidtermo

Os resultados obtidos para o padrão flavonoídico de rutina para as cromatoplacas registradas sob o comprimento de onda de 254 nm estão apresentados na tabela 12. Foi verificado para o teste divisão percentual que a maioria dos modelos de Kernel apresentou excelentes desempenhos na correlação dos atributos de imagem com a respectiva quantificação do analito, pois os valores dos EPMS apresentaram-se consideravelmente baixos. O modelo Kernel RBF não indicou bom desempenho para relacionar os dados de imagem à concentração, cujo valor de EPM foi igual a 0,994. O teste validação cruzada exibiu valores de EPM um pouco maiores, no entanto essa diferença de valor não exibe diferença relevante.

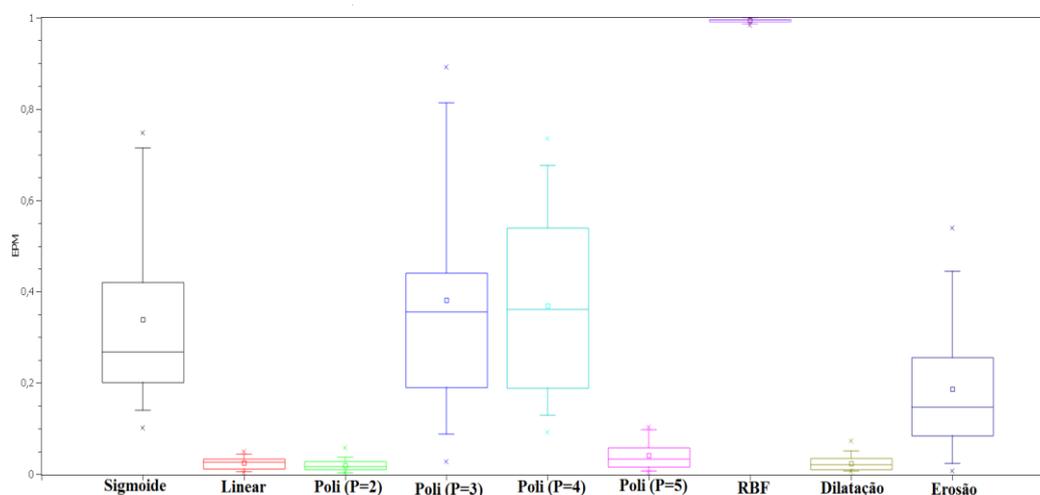
Tabela 12 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 254 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,025	0,024	0,019	0,021	0,033	0,041	0,994	0,023	0,180
Validação cruzada	0,031	0,029	0,024	0,034	0,048	0,059	0,992	0,023	0,022

Fonte: Autoria própria.

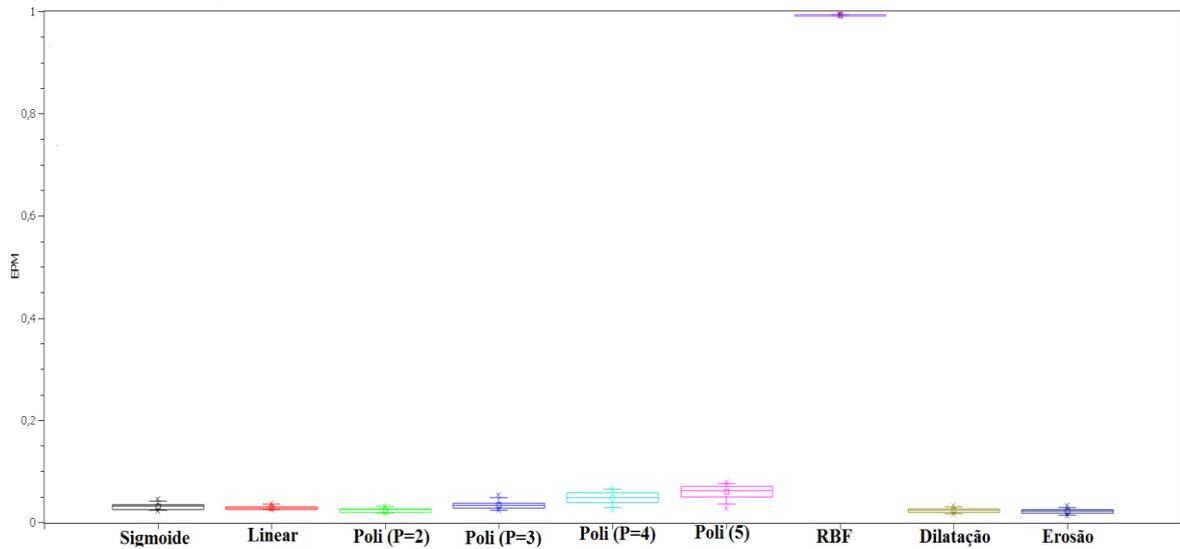
A escolha do melhor teste para desenvolvimento de rede neural artificial (RNA) fica melhor evidenciada ao realizar análise dos gráficos de bloxplots (Figuras 35 e 36). O Boxplot para o teste divisão percentual em 254 nm mostra que os modelos de Kernel sigmoide, polinomial 3, 4 e erosão apresentaram altos indicativos de desvio padrão. Assim, para o teste divisão percentual, esses modelos não são escolhidos como os ideais, devido à difusão de resultados. Ao analisar os resultados obtidos através do teste de validação cruzada, verifica-se que todos os modelos de Kernel analisados apresentaram baixo desvio padrão. Dessa forma, o teste de validação cruzada e os modelos sigmoide, linear, polinomiais, dilatação e erosão apresentam melhores desempenhos para correlação de dados de imagem com a concentração do padrão de rotina para as cromatoplasmas registradas em 254 nm.

Figura 35 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão rotina em 254 nm.



Fonte: Autoria própria.

Figura 36 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão rotina em 254 nm.



Fonte: Autoria própria.

Os resultados provenientes da base de dados dos atributos de imagem gerados pelo Sidtermo das cromatoplacas digitalizadas sob o comprimento de onda de 366 nm estão apresentados na tabela 13. Para ambos os testes aplicados é possível verificar que todos os modelos, excetuando-se a função de base radial, indicaram bons desempenhos na correlação de dados de imagem com a faixa de concentração aplicada.

Tabela 13 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão rotina em 366 nm dos atributos de imagem gerados pelo Sidtermo.

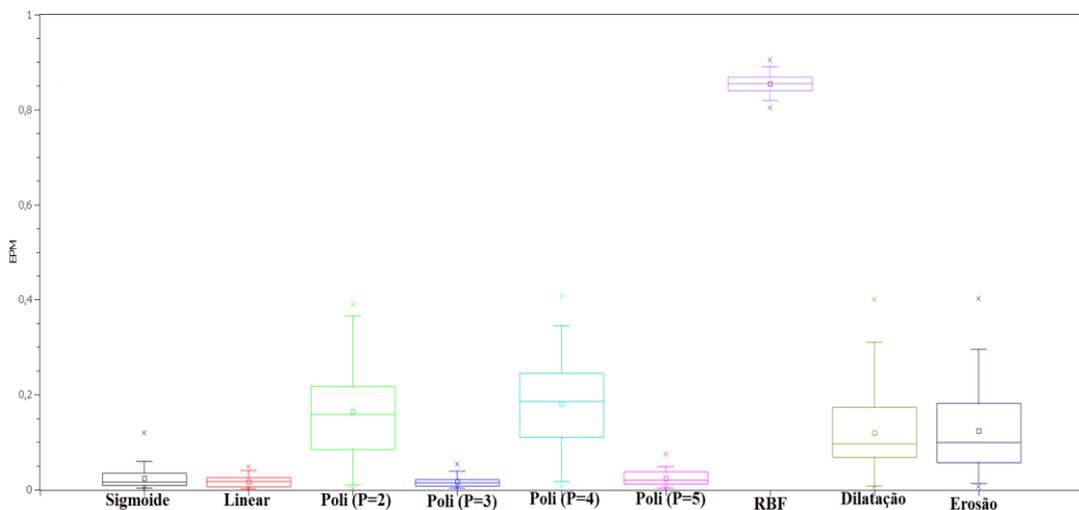
Modelos									
Teste	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,023	0,017	0,016	0,015	0,180	0,023	0,854	0,011	0,012
Validação cruzada	0,032	0,022	0,017	0,025	0,026	0,029	0,852	0,017	0,018

Fonte: Autoria própria.

Para melhor analisar os resultados, os EPMs e o desvio padrão de cada modelo testado estão apresentados nos gráficos de Boxplot, onde a figura 37 apresenta os

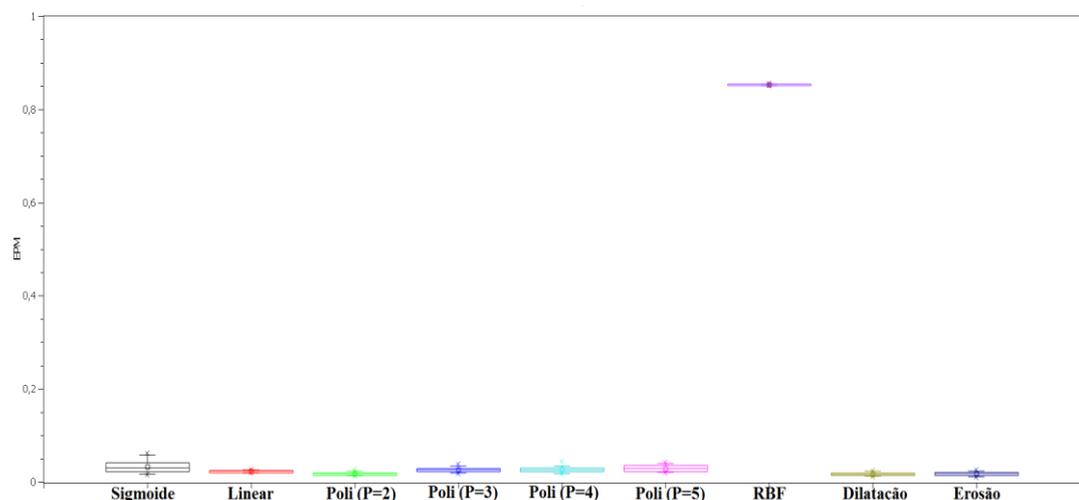
dados obtidos para o teste divisão percentual e a figura 38 apresenta os dados para o teste validação cruzada. Ao analisar o gráfico de Boxplot para o primeiro teste verifica-se que modelos polinomiais 2, 4, dilatação e erosão apresentaram indicativo de desvio padrão consideráveis quando comparados aos demais modelos. Enquanto para o teste validação cruzada, todos os modelos apresentaram baixos desvios de resultados. Assim, o teste validação cruzada e os modelos de Kernel sigmoide, linear, polinomiais, dilatação e erosão são as melhores condições para modelagem de dados do padrão rotina registrado sob o comprimento de onda de 366 nm.

Figura 37 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão rotina em 366 nm.



Fonte: Autoria própria.

Figura 38 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão rotina em 366 nm.



Fonte: Autoria própria.

4.2.4 Ácido cafeico

4.2.4.1 Análise de imagem por ImageJ

Para o teste divisão percentual com relação aos modelos de Kernel estudados para os dados de imagem provenientes do comprimento de onda em 254 nm foi verificado que o modelo linear, dilatação e erosão apresentaram os menores EPM (Tabela 14) indicando maior desempenho para correlação de dados. Para o teste de validação cruzada foi verificado que os modelos dilatação e erosão apresentaram melhor desempenho frente aos demais.

Tabela 14 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 254 nm com atributos de imagem gerados pelo ImageJ.

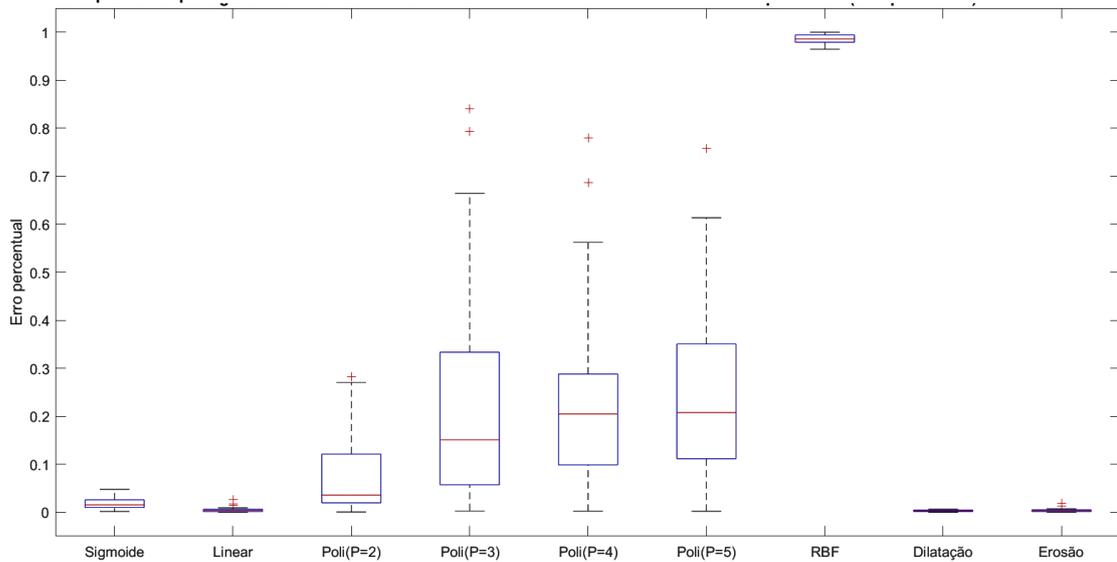
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,038	0,002	0,013	0,022	0,037	0,079	0,977	0,002	0,002
Validação cruzada	0,022	0,013	0,159	0,172	0,204	0,223	0,982	0,005	0,005

Onde: S-Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

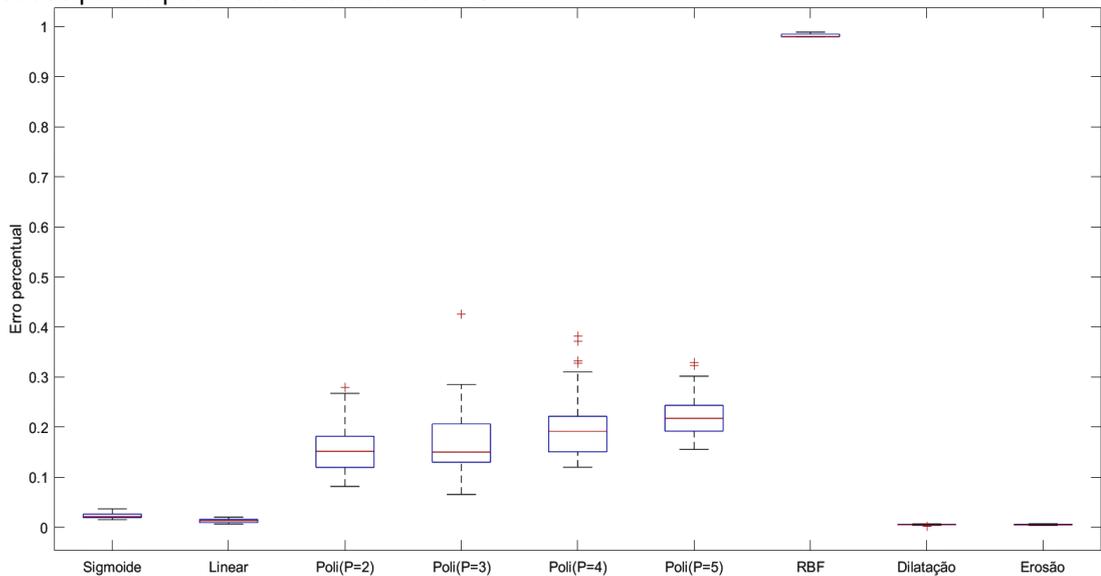
Os gráficos de Boxplot para ambos os testes, divisão percentual e validação cruzada, estão apresentados nas figuras 39 e 40, respectivamente. O gráfico de divisão percentual mostra que o Kernel polinomial 2, 3, 4 e 5 apresentaram os maiores valores de desvio padrão e presença de *outliers*, indicando que esses modelos não apresentaram desempenho satisfatório como os demais. Assim, esses modelos não apresentam de forma fidedigna a correlação de atributos de imagem com a respectiva concentração aplicada.

Figura 39 - Boxplot das imagens tratadas pelo ImageJ para o teste de divisão percentual dos modelos analisados para o padrão rotina em 254 nm.



Fonte: Autoria própria.

Figura 40 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 254 nm.



Fonte: Autoria própria.

Para as imagens registradas em 366 nm foram verificados os menores erros percentuais médios para ambos os testes (divisão percentual e validação cruzada) quando os modelos de kernel sigmoide, linear, dilatação e erosão foram aplicados. Os valores de EPMs para todos os modelos testados estão apresentados na tabela 15.

Tabela 15 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 366 nm com atributos de imagem gerados pelo ImageJ.

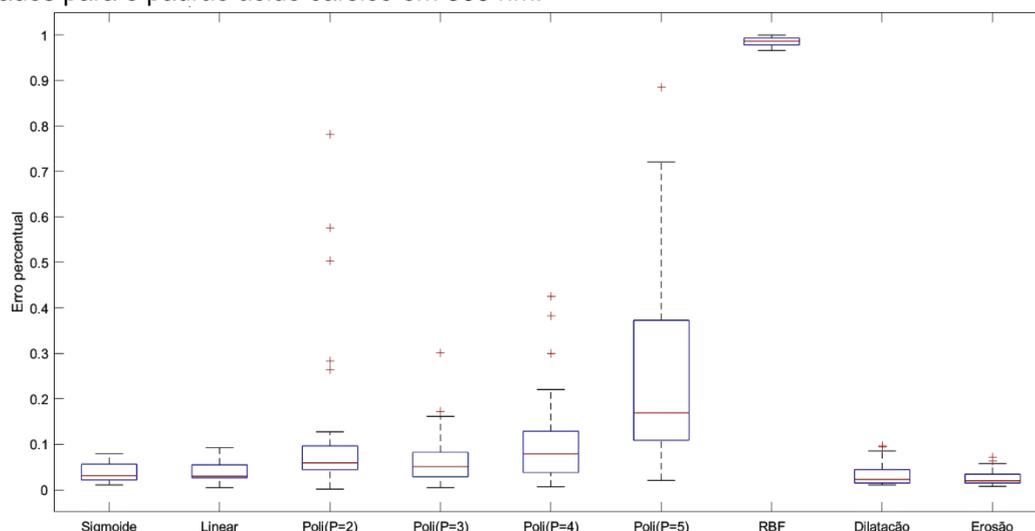
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,015	0,016	0,022	0,069	0,097	0,249	0,985	0,011	0,001
Validação cruzada	0,024	0,036	0,085	0,073	0,106	0,166	0,981	0,018	0,021

Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

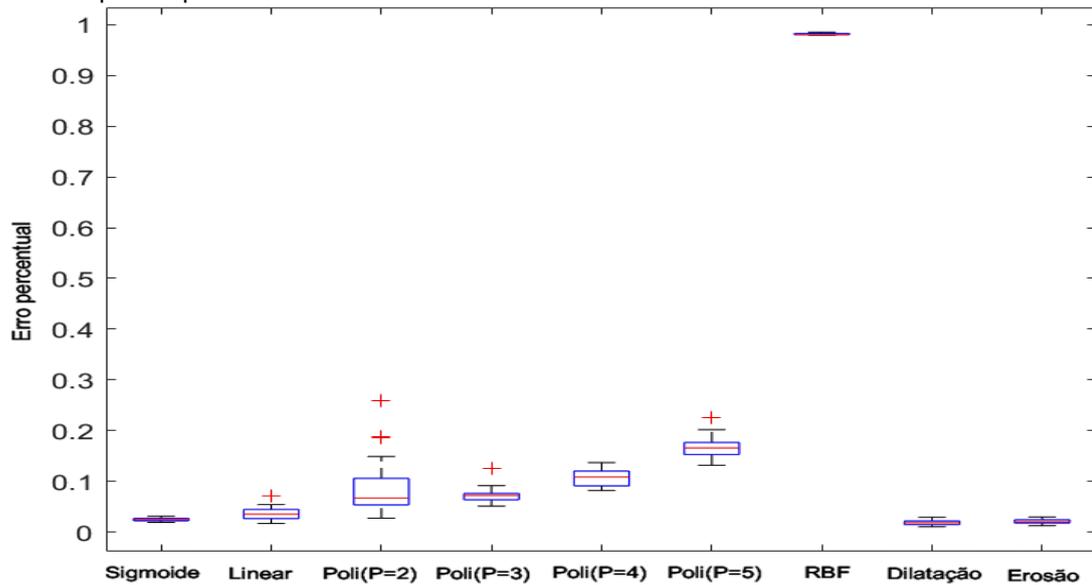
Para melhor visualização e interpretação dos resultados, o gráfico de Boxplot para os modelos divisão percentual e validação cruzada estão apresentados respectivamente nas figuras 41 e 42. Ao analisar a figura 41 é possível verificar que os modelos polinomiais apresentaram os maiores desvio padrão quando comparado aos demais modelos de Kernel. Além disso, foram os modelos que mais apresentaram curvas fora do padrão quando comparadas umas com as outras. Para o teste validação cruzada foi verificado comportamento semelhante com relação parâmetro estatístico e presença de sinais *outliers* (+).

Figura 41 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão ácido cafeico em 366 nm.



Fonte: Autoria própria.

Figura 42 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 366 nm.



Fonte: Autoria própria.

4.2.4.2 Análise de imagem Sidtermo

Os resultados dos EPMs obtidos da base de dados provenientes do software Sidtermo para o padrão de ácido cafeico estão apresentados na tabela 16. Os dois tipos de teste aplicados para todos os modelos analisados indicaram baixos EPMs, com exceção de Kernel RBF, cujos valores correspondem a 0,999 e 0,998 para o teste divisão percentual e validação cruzada, respectivamente.

Tabela 16 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 254 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,010	0,013	0,009	0,011	0,013	0,014	0,999	0,009	0,008
Validação cruzada	0,014	0,012	0,010	0,015	0,023	0,021	0,998	0,011	0,011

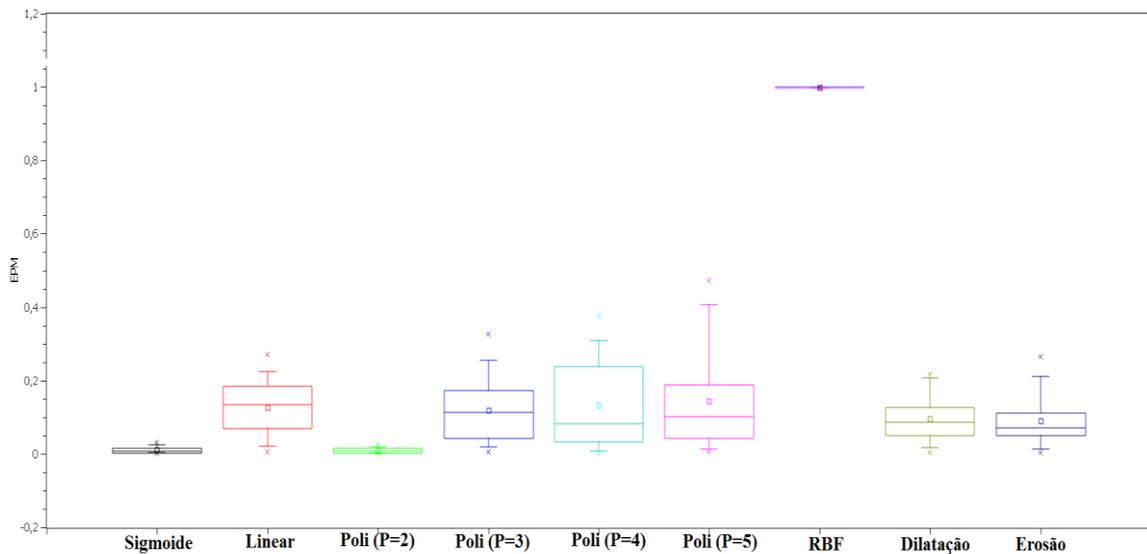
Fonte: Autoria própria.

Os gráficos de bloxplot, para os modelos analisados com os EPMs obtidos a partir da correlação dos atributos de imagem gerados pelo Sidtermo das cromatoplasmas registradas sob o comprimento de onda de 254 nm, estão apresentados nas figuras 43 e 44. Foi possível constatar que apenas os modelos sigmoide, polinomial de grau 2 e RBF apresentaram baixo desvio padrão, os demais modelos apresentaram discrepância de resultados verificada tanto pelo desvio padrão quanto presença de sinais *outliers* (+). Assim, os modelos sigmoide e polinomial 2 são escolhidos como ideais para desenvolvimento de RNA quando o teste de divisão percentual for aplicado.

Ao avaliar o teste validação cruzada, verifica-se que todos os modelos analisados apontaram desvio padrão consideravelmente baixo. Dessa maneira, para o segundo teste todos os modelos são ideais para correlação de dados, excetuando-se a função de base radial devido ao elevado EPM que o modelo apresentou.

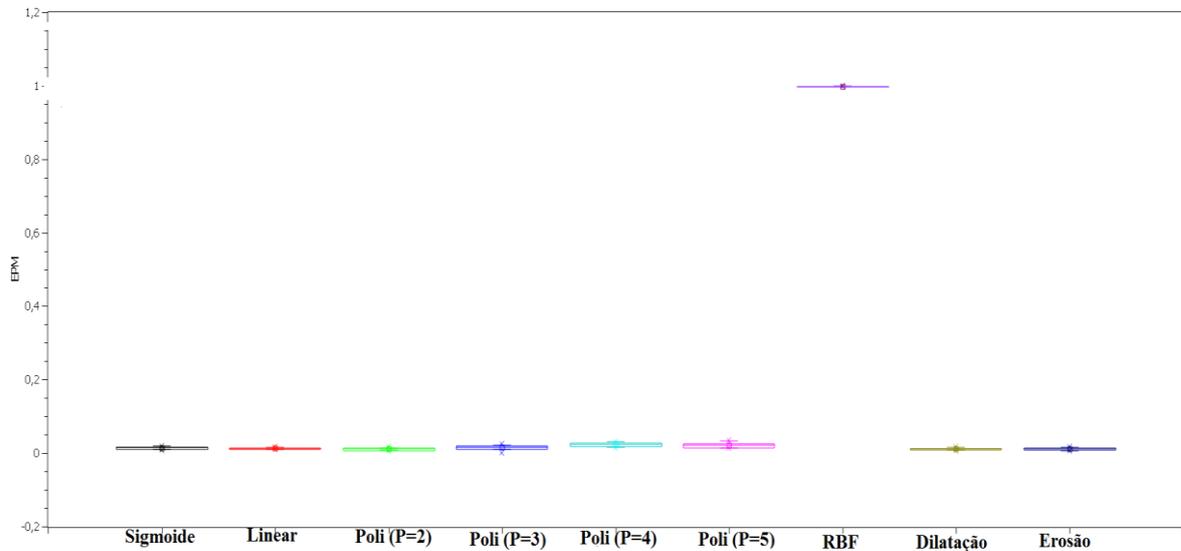
Para o padrão ácido cafeico registrado sob o comprimento de onda de 254 nm fica estabelecido como as melhores condições de análise o teste de validação cruzada dos modelos sigmoide, linear, polinomiais dilatação e erosão.

Figura 43 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido cafeico em 254 nm.



Fonte: Autoria própria.

Figura 44 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 254 nm.



Fonte: Autoria própria.

A tabela 17 mostra os valores de EPM para os modelos analisados para modelagem de dados provenientes do Sidtermo. O comportamento de ambos os testes foram os mesmos descritos para o comprimento de onda de 254 nm.

Tabela 17 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão ácido cafeico em 366 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,013	0,010	0,007	0,008	0,012	0,015	1,000	0,008	0,007
Validação cruzada	0,015	0,013	0,010	0,013	0,017	0,017	1,000	0,011	0,011

Fonte: Autoria própria.

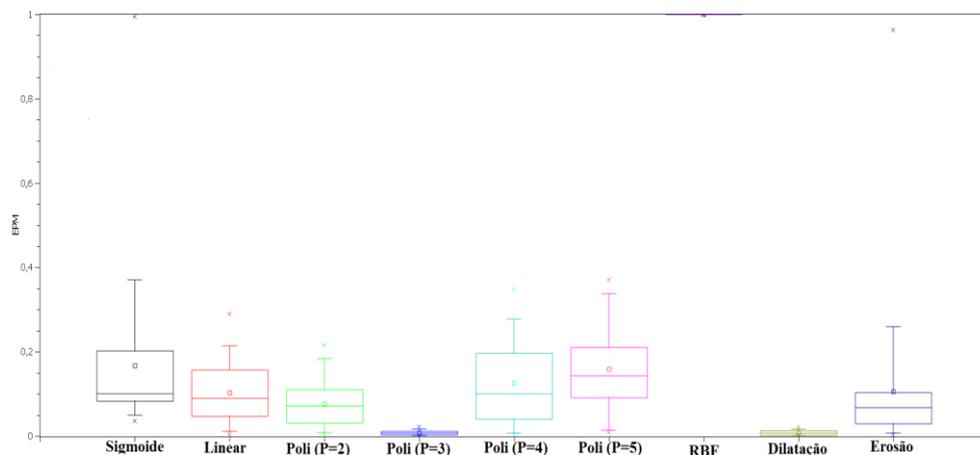
O gráfico de bloxplot para o teste divisão percentual (Figura 45) mostra que apenas os modelos Kernel polinomial de grau 3, RBF e dilatação apresentaram baixa difusão de resultados. Assim, os modelos que evidenciam melhor desempenho justificado pelo baixo EPM e desvio padrão são os modelos polinomial 3 e dilatação.

O gráfico de Boxplot para o teste validação cruzada (Figura 46) mostra que todos os modelos apresentaram difusão de dados significativamente baixa. Além disso, não foi detectada a presença de sinais *outliers* (+).

Portanto, os modelos sigmoide, linear, polinomiais, dilatação e erosão apresentam-se como bons modeladores de dados devido aos baixos erros percentuais e baixo desvio padrão.

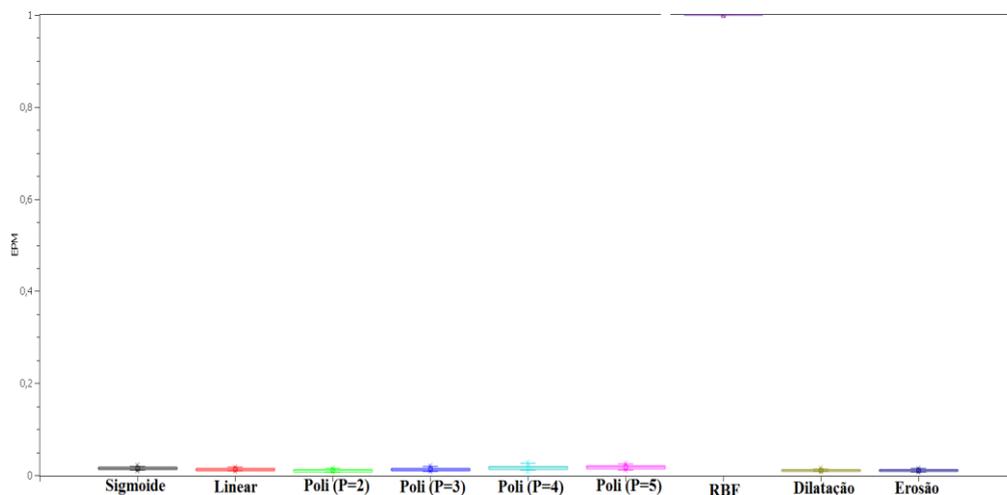
Para o padrão de ácido cafeico em 366 nm, considera-se que o teste de validação cruzada com os modelos sigmoide, linear, polinomiais, dilatação e erosão representam as melhores condições para correlação dos atributos de imagem gerados pelo Sidtermo com a respectiva faixa de concentração empregada.

Figura 45 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão ácido cafeico em 366 nm.



Fonte: Autoria própria.

Figura 46 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão ácido cafeico em 366 nm.



Fonte: Autoria própria.

4.2.5 Catequina

4.2.5.1 Análise de imagem ImageJ

Com relação aos testes realizados, para o padrão de catequina registrado sob o comprimento de onda de 254 nm foi verificado que para o teste de divisão percentual o EPM foi menor para os modelos linear, dilatação e erosão, cujos valores estão apresentados na tabela 18. O mesmo desempenho dos modelos testados foi verificado para o teste de validação cruzada, onde os valores dos erros percentuais médios foram iguais a 0,004.

Tabela 18 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 254 nm com atributos de imagem gerados pelo ImageJ.

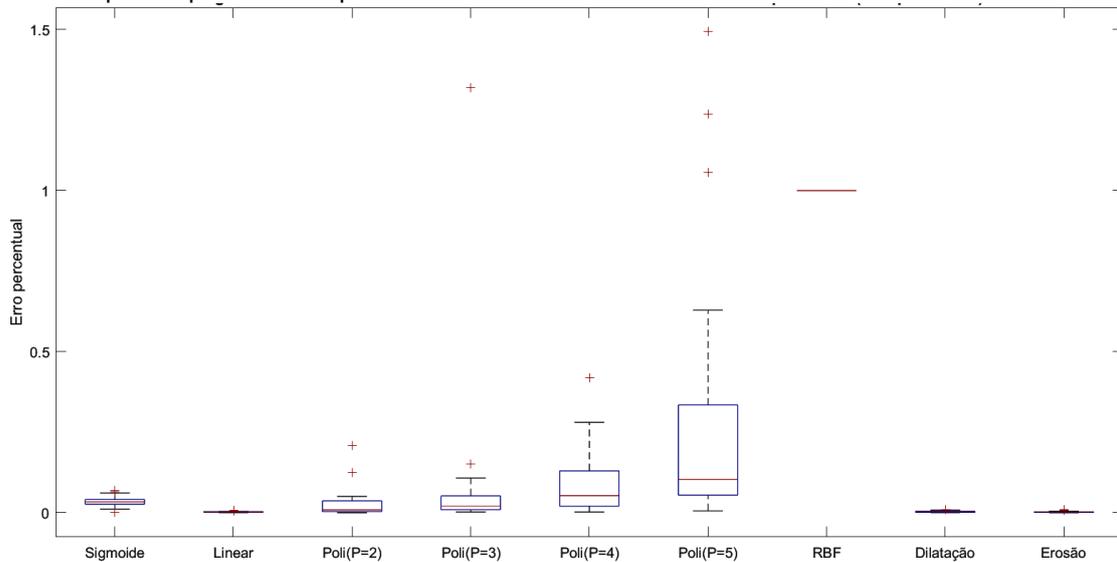
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,002	0,025	0,077	0,093	0,276	0,079	1,000	0,002	0,002
Validação cruzada	0,036	0,004	0,002	0,035	0,103	0,248	1,000	0,004	0,004

Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

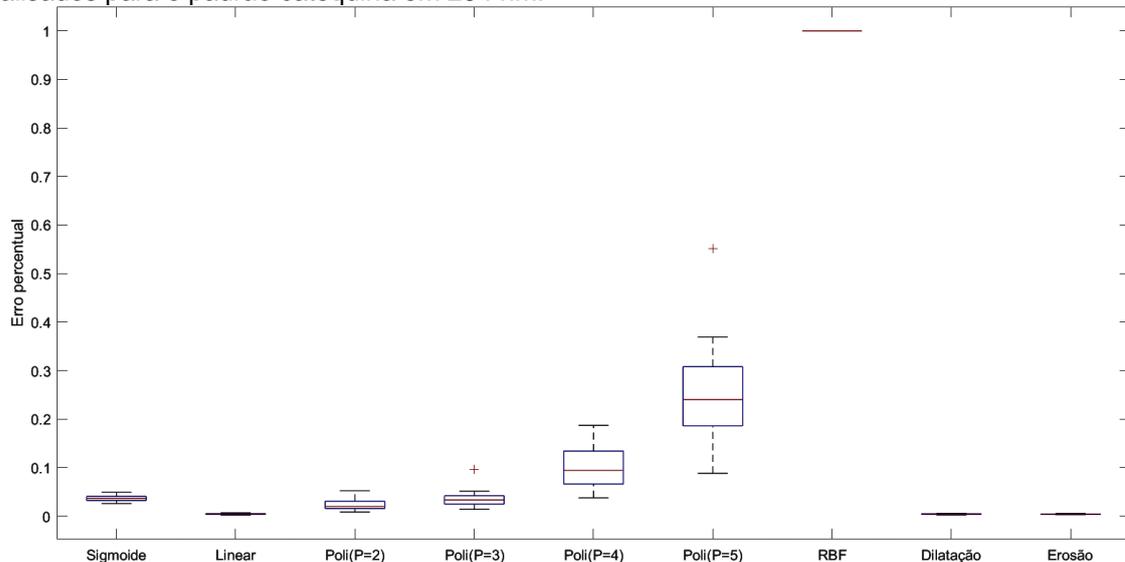
Os gráficos de Boxplot para os testes realizados com as cromatoplasmas registradas sob o comprimento de onda de 254 nm estão apresentados nas figuras 47 e 48. Para o teste de divisão percentual é possível visualizar que os maiores valores de desvio padrão foram para os modelos de Kernel polinomiais. Além disso, foi observada a presença de variação atípica indicada através do sinal de resíduo (+). O mesmo comportamento de desvio padrão foi verificado para o teste de validação cruzada. No entanto, esse tipo de teste indicou apenas a presença de dois sinais outliers (+), um sinal para o polinômio de grau 3 com valor aproximadamente de 0,1 e outro no polinomial de grau 5, cujo valor foi superior a 0,5.

Figura 47 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão catequina em 254 nm.



Fonte: Autoria própria.

Figura 48 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão catequina em 254 nm.



Fonte: Autoria própria.

Ao realizar a análise das imagens em 366 nm para o padrão catequina foi observado bom desempenho dos modelos de kernel linear, dilatação e erosão quando o teste de divisão percentual foi aplicado. Enquanto o teste validação cruzada apresentou desempenho os modelos sigmoide, linear, dilatação e erosão. A tabela 19 mostra os valores de EPM para os testes e modelos realizados.

Tabela 19 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 366 nm com atributos de imagem gerados pelo ImageJ.

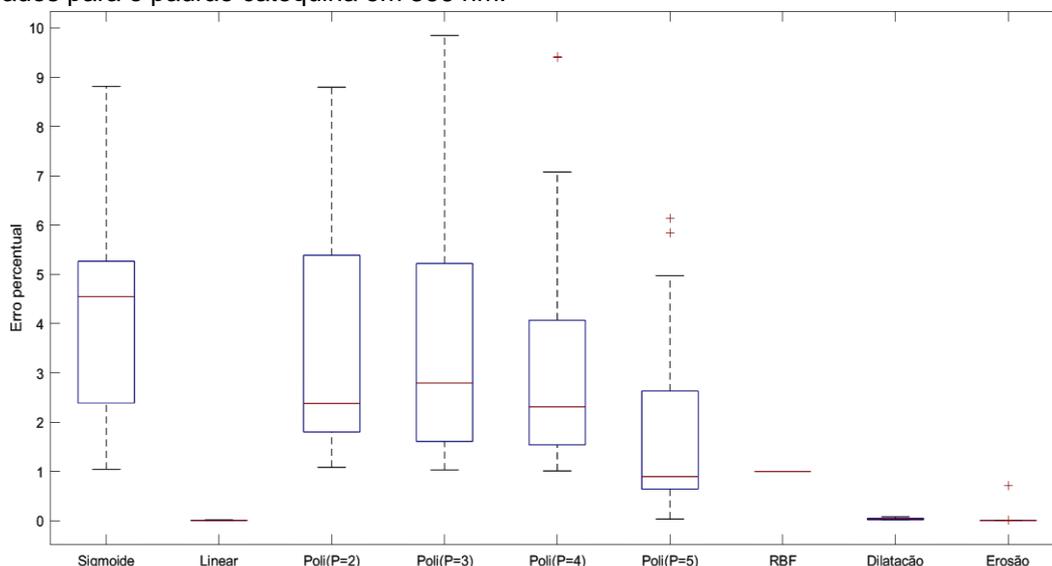
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	3,638	0,002	4,156	1,265	1,401	1,304	1,000	0,002	0,002
Validação cruzada	0,039	0,004	1,215	211,350	3,786	4,652	1,000	0,003	0,003

Onde: S - Sigmoide; P – polinomial (grau); D – dilatação; E – erosão.

Fonte: Autoria própria.

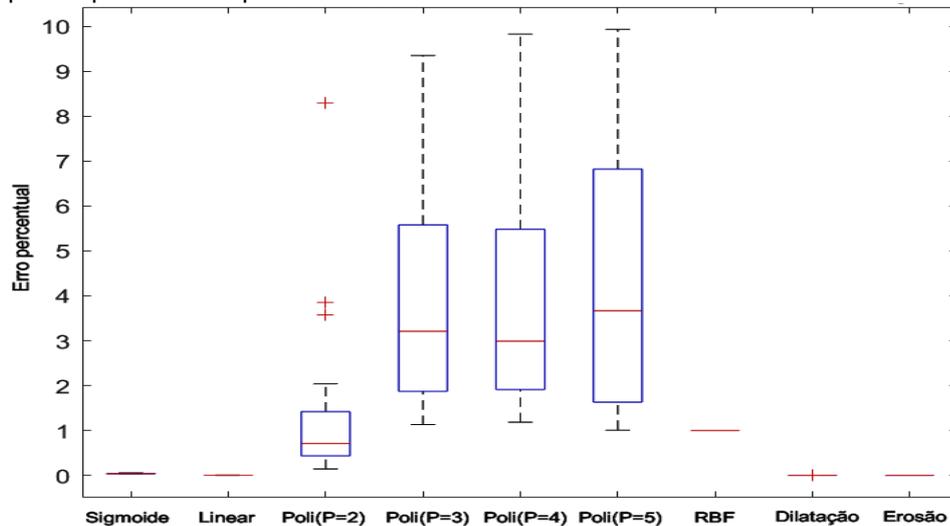
A figura 49 confirma que os modelos sigmoide e polinomial não são ideais para correlação dos dados de imagem com a concentração do analito por ter apresentado alto valor de EPM e desvio padrão quando o teste de divisão percentual foi aplicado. Para o teste de validação cruzada (Figura 50) o gráfico de Boxplot mostra que os modelos polinomiais 3, 4 e 5 apresentaram desvio padrão significativos, além de terem apresentado alto valor de EPM. Portanto, não representam os melhores modelos de Kernel para modelagem de dados.

Figura 49 - Boxplot das imagens tratadas pelo ImageJ para o teste divisão percentual dos modelos analisados para o padrão catequina em 366 nm.



Fonte: Autoria própria.

Figura 50 - Boxplot das imagens tratadas pelo ImageJ para o teste validação cruzada dos modelos analisados para o padrão catequina em 366 nm.



Fonte: Autoria própria.

4.2.5.2 Análise de imagem por Sidtermo

Os resultados dos EPMs para o padrão de catequina (digitalização em 254 nm) considerando a correlação dos atributos de imagem gerados pelo Sidtermo estão apresentados na tabela 20. Segundo os valores dos EPMs, para o teste divisão percentual, os modelos de Kernel que apresentaram melhor desempenho de associação entre os atributos de imagem com a faixa de concentração aplicada foram os polinômios de graus 3, 4 e 5 e o modelo linear, cujos valores foram significativamente baixos. A função de base radial foi o modelo que indicou maior valor correspondente a 0,997. Enquanto o teste de validação cruzada revelou desempenho de correlação para todos os modelos analisados, excetuando-se o modelo de RBF.

Tabela 20 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 254 nm dos atributos de imagem gerados pelo Sidtermo.

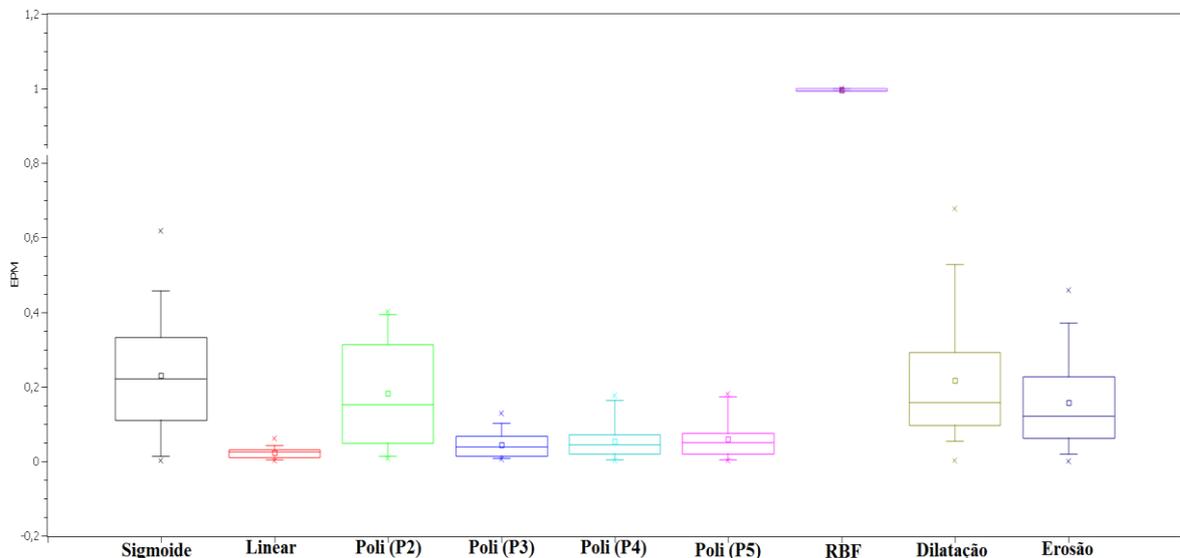
Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,254	0,022	0,181	0,044	0,053	0,059	0,997	0,216	0,157
Validação cruzada	0,028	0,028	0,024	0,053	0,075	0,094	0,997	0,021	0,022

Fonte: Autoria própria.

Os gráficos de Boxplots para as cromatoplasmas registradas em 254 nm estão apresentados nas figuras 51 e 52. Para o teste divisão percentual, os modelos que indicaram melhor desempenho para correlação de dados apresentaram baixo desvio padrão. O modelo de Kernel RBF também indicou baixo desvio padrão, no entanto devido ao alto EPM esse modelo não é considerado o ideal. Observa-se a presença de um resíduo discrepante para os modelos sigmoide, dilatação e erosão com valores aproximados de 0,6; 0,7 e 0,5, respectivamente. Esse resíduo corresponde aos *outliers*, que podem ocorrer segundo as causas descritas anteriormente. Para o teste de validação cruzada, foi verificado que todos os modelos apresentaram baixo desvio padrão e nenhum sinal de resíduo foi detectado.

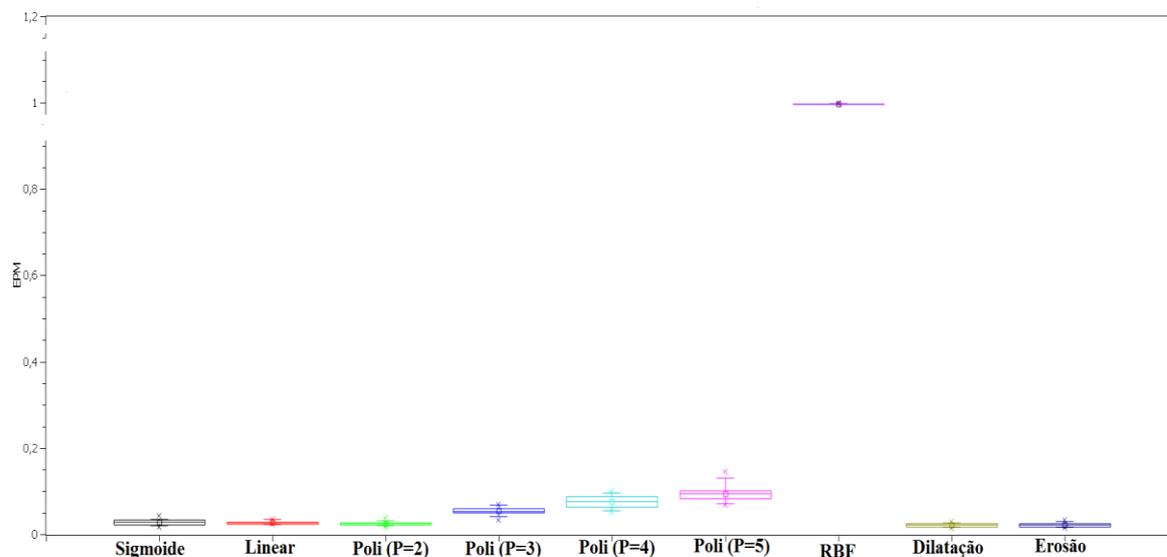
Para o padrão de catequina sob o comprimento de onda de 254 nm fica estabelecido como as melhores condições para o desenvolvimento de RNA o teste de validação cruzada para os modelos sigmoide, linear, polinomiais, dilatação e erosão. O software Sidtermo também foi escolhido como a melhor ferramenta de imagem, devido ao tempo de análise ser consideravelmente menor, por ter menos interferência do operador e por gerar resultados tão precisos quanto o ImageJ.

Figura 51 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão catequina em 254 nm.



Fonte: Autoria própria.

Figura 52 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão catequina em 254 nm.



Fonte: Autoria própria.

Os resultados dos EPMs para o padrão de catequina em 366 nm considerando a correlação dos atributos de imagem gerados pelo Sidtermo estão apresentados na tabela 21. Os modelos que indicaram melhor desempenho foram o sigmoide, polinomial 2, 4 e 5 para o teste de divisão percentual. O teste de validação cruzada evidenciou bom desempenho para todos os modelos testados, com exceção do modelo de RBF.

Tabela 21 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 366 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,024	0,231	0,019	0,297	0,052	0,050	1,000	0,165	0,198
Validação cruzada	0,027	0,027	0,021	0,047	0,071	0,087	1,000	0,020	0,022

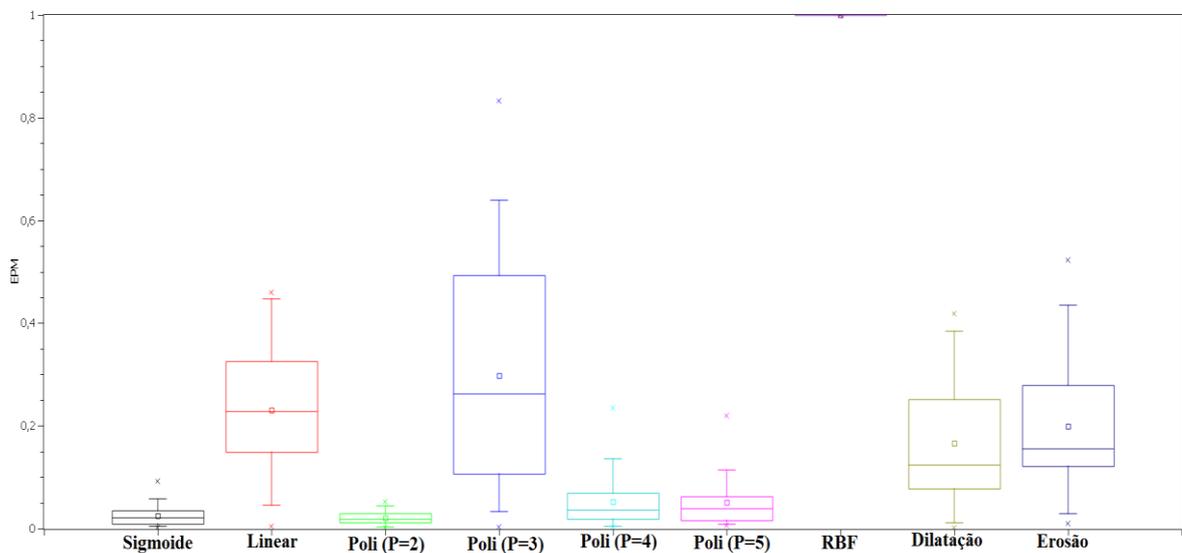
Fonte: Autoria própria.

Os Boxplots para o teste divisão percentual e validação cruzada estão apresentados nas figuras 53 e 54, respectivamente. Os modelos que indicaram melhor

desempenho apresentaram menor desvio padrão para ambos os testes. O teste de divisão percentual para os modelos linear, polinomial 3, dilatação e erosão indicou a presença de sinal de resíduo.

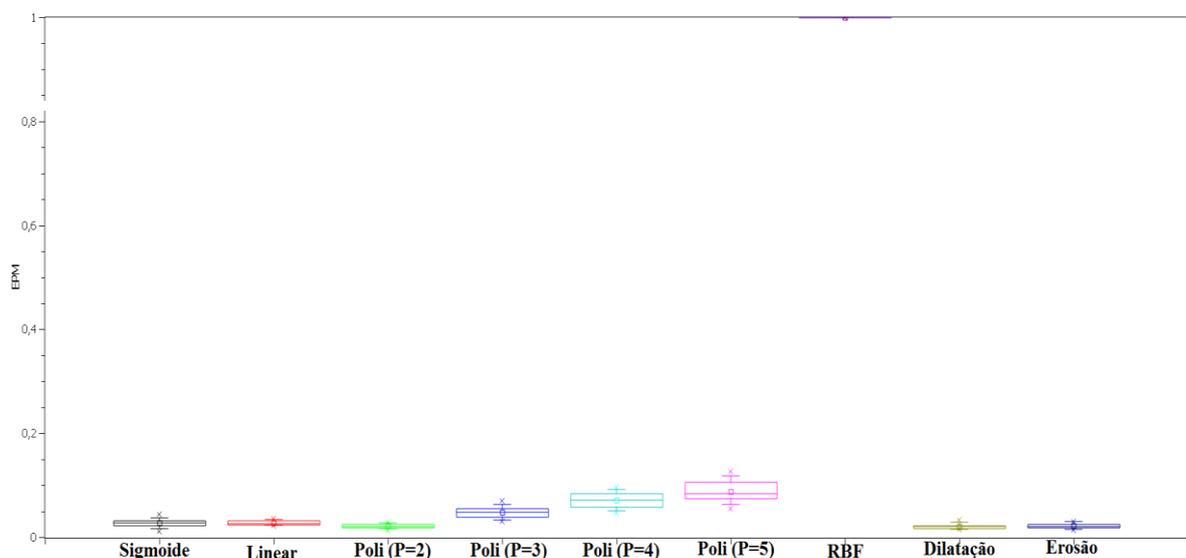
As condições mais adequadas para o padrão catequina em 366 nm correspondem ao teste de validação cruzada e os modelos sigmoide, linear, polinomiais, dilatação e erosão. A ferramenta de imagem Sidtermo foi escolhida como a mais ideal para tratamento das imagens.

Figura 53 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para o padrão catequina em 366 nm.



Fonte: Autoria própria.

Figura 54 - Boxplot das imagens tratadas pelo Sidtermo para o teste validação cruzada dos modelos analisados para o padrão catequina em 366 nm.



Fonte: Autoria própria.

4.2.6 *Schinus terebinthifolius*

Os resultados das configurações de redes neurais artificiais obtidos dos padrões de polifenóis foram essenciais para o desenvolvimento dos testes nas espécies vegetais. Como visto anteriormente, a ferramenta de imagem escolhida para posteriores análises foi o Sidtermo, devido a precisão de resultados verificada através dos baixos EPMS, rapidez de execução de análise e baixa interferência do operador na manipulação das imagens. As análises realizadas pelo ImageJ necessitavam de tratamento prévio para posterior análise e demandavam muito tempo para obtenção da base de dados.

Os resultados obtidos para *S. terebinthifolius* com os dados gerados pelo Sidtermo para as imagens registradas sob o comprimento de onda de 254 nm estão apresentados na tabela 22. Para o teste de divisão percentual, foi verificado que os modelos sigmoide, linear, polinomiais de todos os graus analisados apresentaram excelentes desempenhos para correlacionar os atributos de imagem com o teor de ácido gálico calculado na cromatoplaça através da curva de calibração obtida por cromatografia líquida de alta eficiência. Assim, os teores de ácido gálico empregados na cromatoplaça para as amostras de aroeira na faixa de concentração de 30-40 µg/mL foi de 0,0013; 0,0014; 0,0015; 0,0016 e 0,0017 g%.

Para o teste de validação cruzada, não foi possível obter resultados, pois quando a base de dados foi colocada para gerar resultados no Octave o sistema apresentou erros. Isso pode ter sido possivelmente devido ao tamanho da base de dados ter sido insuficiente.

Os resultados apontam êxito na correlação de dados de imagem com o teor, sendo possível quantificar ácido gálico em aroeira através de sistemas de análise imagem.

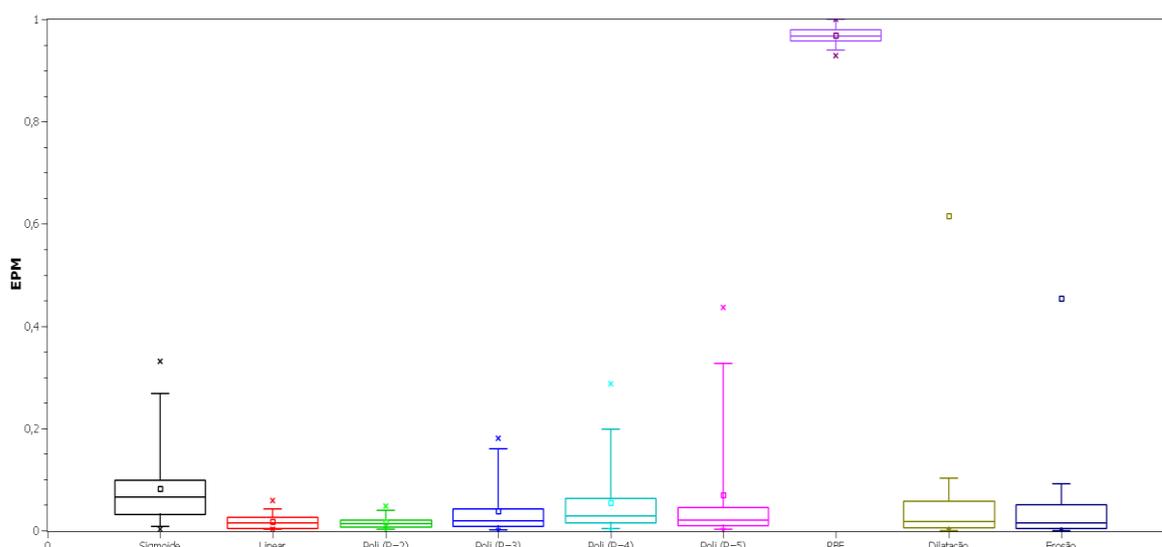
Tabela 22 - Valores de erro percentual médio (EPM) para cada modelo analisado para *S. terebinthifolius* em 254 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,082	0,018	0,016	0,038	0,055	0,070	0,969	0,616	0,454

Fonte: Autoria própria.

O Boxplot para os resultados da correlação dos dados de imagem com o respectivo teor de ácido gálico presente na espécie para as cromatoplasmas registradas sob o comprimento de onda de 254 nm está apresentado na figura 55. De acordo com a figura é possível verificar que os modelos de Kernel linear e polinomial de grau 2 evidenciaram a menor discrepância de resultados quando o sistema analisou por 30 vezes a base de dados. Dentre os modelos analisados o Kernel sigmoide foi o que indicou maior desvio padrão. Dessa forma, considerando os EPMs e a variância de resultados nas análises, os modelos de Kernel escolhidos para modelagem de dados do marcador ácido gálico são o linear e polinomial de grau 2 para as imagem registradas sob o comprimento de onda de 254 nm.

Figura 55 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de *S. terebinthifolius* em 254 nm.



Fonte: Autoria própria.

Os resultados obtidos para as cromatoplasmas digitalizadas sob o comprimento de onda de 366 nm estão apresentados na tabela 23. Verifica-se que para o teste de divisão percentual, todos os modelos de Kernel apresentaram bons desempenhos para modelagem de dados, com exceção da função de base radial. Para o teste validação cruzada o mesmo erro de análise foi verificado como descrito anteriormente para as placas cromatográficas registradas sob o comprimento de onda de 254 nm.

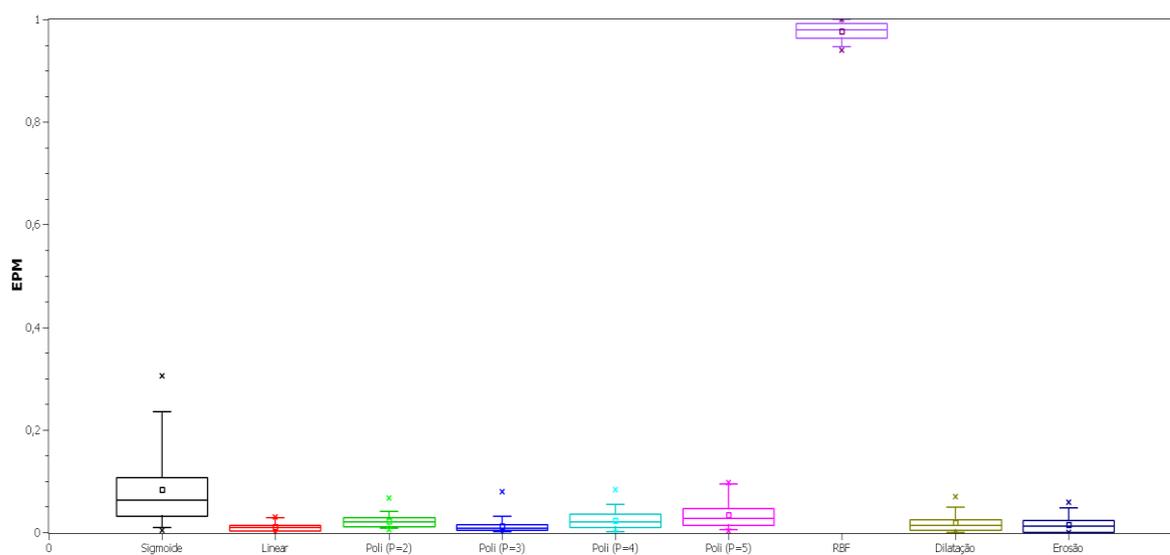
Tabela 23 - Valores de erro percentual médio (EPM) para cada modelo analisado para *S. terebinthifolius* em 366 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,082	0,010	0,022	0,012	0,023	0,034	0,977	0,019	0,015

Fonte: Autoria própria.

O gráfico de Boxplot para o teste anteriormente descrito está apresentado na figura 56. É possível verificar que o modelo de Kernel sigmoide apresentou maior discrepância de resultados, seguido dos modelos polinomial de graus 4, 5 e função de radial. Os demais modelos indicaram baixo desvio padrão. Assim, considerando os resultados apresentados na tabela e o gráfico de Boxplot, ficam estabelecidos os modelos com melhores desempenho para correlação dos atributos de imagem com o respectivo teor de ácido gálico para as cromatoplasmas digitalizadas sob o comprimento de onda de 366 nm os modelos de Kernel linear, polinomiais de graus 2 e 3, dilatação e erosão.

Figura 56 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de *S. terebinthifolius* em 366 nm.



Fonte: Autoria própria.

Além da análise da espécie pelo Octave[®] outra ferramenta de modelagem de dados foi aplicada, o Weka, com a finalidade de verificar mais um software que permita

a correlação de dados matemáticos provenientes das imagens com a concentração. Dois parâmetros estatísticos são considerados para análise dos resultados, o erro quadrático médio (EQM) e o coeficiente de correlação. Para as imagens registradas sob o comprimento de onda de 254 nm, a média das 30 leituras realizadas tanto para o EQM quanto para o coeficiente de correlação foi de 87,45 e 0,6, respectivamente. O valor para o EQM foi muito alto, quanto menor o valor desse parâmetro mais os resultados experimentais se aproximam do real. Enquanto o coeficiente de determinação apresentou-se baixo, indicando que houve baixa associação dos atributos de imagem com a concentração do analito de interesse. Para as imagens digitalizadas sob o comprimento de onda de 366 nm, o valor do EQM foi de 110,56 e o coeficiente de determinação foi igual a 0,35. Logo, o Weka ou as condições empregadas não são consideradas ideais para correlacionar dados de imagem.

4.2.7 *Anadenanthera colubrina*

Os resultados obtidos para *A. colubrina*, da base de dados provenientes do Sidtermo, para as imagens registradas sob o comprimento de onda de 254 nm estão apresentados na tabela 24. Como apresentado na tabela para o teste divisão percentual a espécie de *A. colubrina* alguns modelos apresentaram bom desempenho para modelagem de dados. Para o teste de divisão percentual, os modelos de Kernel que apresentaram os menores erros percentuais médios foram o linear, os modelos polinomiais, dilatação e erosão. Os resultados indicam que é possível correlacionar os atributos de imagem com o teor de catequina calculado na cromatoplaça através da curva de calibração obtida por cromatografia em camada delgada de alta eficiência. Assim, os teores de catequina empregados na cromatoplaça para as amostras de angico na faixa de concentração de 300-400 µg/mL foram de 0,300; 0,325; 0,350; 0,375 e 0,400 g%.

Quando o teste de validação cruzada foi empregado, verificou-se o mesmo erro de análise encontrado para a espécie de *S. terebithinifolia* não sendo possível realizar o teste.

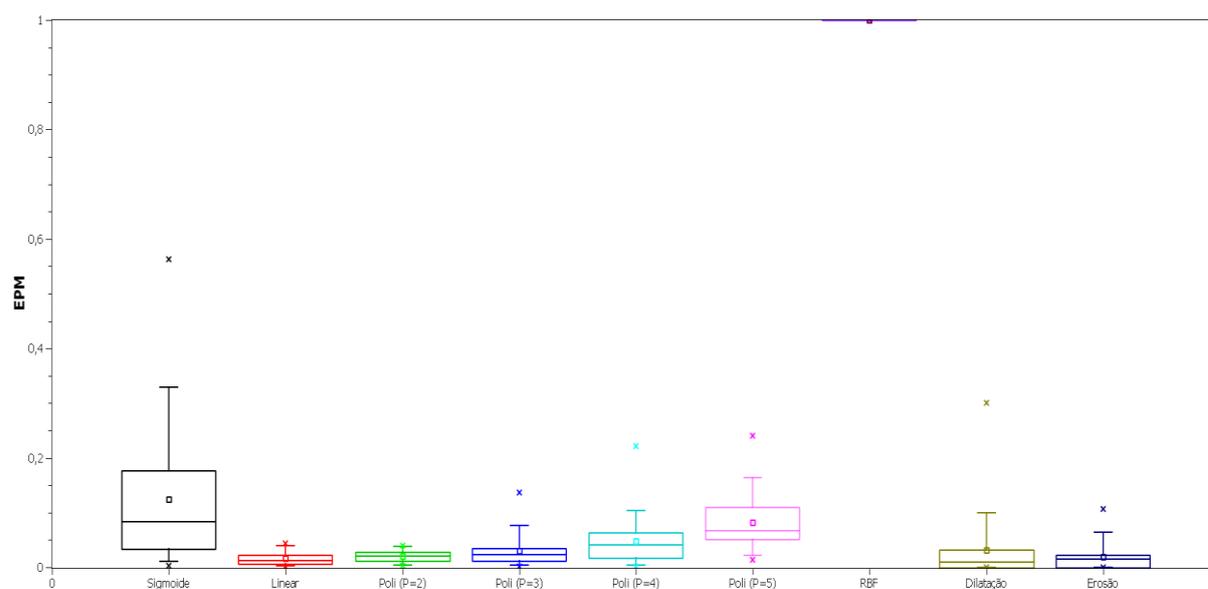
Tabela 24 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 254 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,123	0,014	0,019	0,029	0,041	0,058	0,999	0,018	0,015

Fonte: Autoria própria.

O Boxplot para os resultados da correlação dos dados de imagem do marcador catequina presente na espécie com o respectivo teor para o teste de divisão percentual, com a placa cromatográfica digitalizada em 254 nm está apresentado na figura 57. Verifica-se que o modelo sigmoide apresentou desvio padrão consideravelmente elevado, não sendo, portanto, o modelo mais ideal para correlação de dados. Os modelos de Kernel polinomiais de graus 4 e 5 também evidenciaram discrepância de resultados maior quando comparados aos demais modelos. Dessa forma, ficam estabelecidos os melhores modelos para modelagem de dados, considerando os resultados apontados na tabela 24 e o gráfico, os modelos linear, polinomiais de graus 2 e 3, dilatação e erosão.

Figura 57 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de *A. colubrina* em 254 nm.



Fonte: Autoria própria.

Os resultados obtidos para as imagens digitalizadas sob o comprimento de onda de 366 nm estão apresentados na tabela 25. Verifica-se para o teste divisão percentual, através dos valores de EPM, que os modelos sigmoide, linear, polinomiais de todos os graus analisados, dilatação e erosão apontaram os menores valores de EPMs frente ao modelo de função de base radial. Assim, para correlação dos atributos de imagem do marcador catequina na espécie *A. colubrina* com o respectivo teor aplicado na cromatoplaça ficam estabelecidos os modelos de Kernel sigmoide, linear, polinomiais, dilatação e erosão.

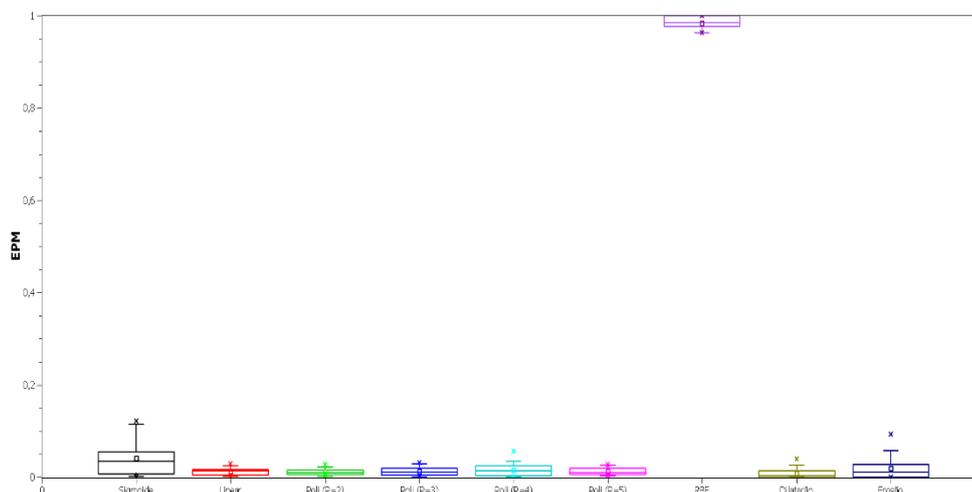
Tabela 25 - Valores de erro percentual médio (EPM) para cada modelo analisado para o padrão catequina em 366 nm dos atributos de imagem gerados pelo Sidtermo.

Teste	Modelos								
	S	Linear	P(2)	P(3)	P(4)	P(5)	RBF	D	E
Divisão percentual	0,055	0,012	0,009	0,013	0,011	0,012	0,984	0,082	0,080

Fonte: Autoria própria.

A figura 58 apresenta o gráfico de Boxplot para os modelos analisados para as cromatoplaças registradas sob o comprimento de onda de 366 nm para os testes de divisão percentual. De acordo com o gráfico e com os resultados apontados na tabela 25, os modelos que são considerados com melhor desempenho e maior precisão de dados são o linear, polinomiais de todos os graus, dilatação e erosão.

Figura 58 - Boxplot das imagens tratadas pelo Sidtermo para o teste divisão percentual dos modelos analisados para a espécie de *A. colubrina* em 366 nm.



Fonte: Autoria própria.

Para a espécie também foram obtidos os valores dos parâmetros dos EQM e coeficiente de correlação através do estudo de regressão realizado através do Weka. A média do EQM e do coeficiente de correlação para as imagens registradas sob o comprimento de onda de 254nm foram iguais a 97,53 e 0,46. Enquanto as imagens digitalizadas sob o comprimento de onda de 366 nm, mostraram o valor de 76,23 para o EQM e 0,72 para o coeficiente de determinação. Dessa forma, diante dos resultados obtidos é possível afirmar que a modelagem de dados pelo Weka não constitui uma alternativa favorável para correlação dos atributos de imagem com o teor de catequina apresentado na espécie.

4.3 ANÁLISE QUANTITATIVA DE POLIFENOIS POR CLAE-DAD

Com a finalidade de comparar os resultados do estudo realizado por imagem com a técnica considerada padrão ouro para quantificação de fitoconstituintes foi realizada a análise por cromatografia líquida de alta eficiência para verificar a relação linear da faixa de concentração aplicada para cada analito. Além disso, a obtenção da curva de calibração por CLAE foi necessária para determinar o teor presente nas cromatoplasmas dos marcadores químicos de ácido gálico e catequina para *S. terebinthifolius* e *A. colubrina*, respectivamente.

Os valores dos coeficientes linear e angular foram obtidos pelo método de mínimos quadrados, com a finalidade de garantir que a reta obtida é aquela para a qual se tem as menores distâncias entre os valores observados e a própria reta (CALLEGARI-JACQUES, 2003). Os valores dos parâmetros estatísticos para estabelecer a equação de regressão e os Limites de detecção (LOQ) e quantificação (LOD) estão apresentados na tabela 26.

Tabela 26 - Avaliação estatística da resposta de linearidade de soluções dos padrões polifenólicos e amostras de de *S. terebinthifolius* e *A. colubrina*.

Analito	Faixa de		a	b	Valor-p	LD	LQ
	concentração (µg/mL)						
Ácido gálico	20-30		1,33	2,30	0,030	2,36	7,15
Quercetina	7,5-17,5		1,46	-3,15	0,003	1,97	5,99
Rutina	15-27		0,93	-8,70	1,720	1,21	3,67
Ácido cafeico	25-35		0,93	-6,18	3,930	2,87	8,69
Catequina	12-20		0,14	-0,19	0,005	1,27	3,85
<i>S. terebinthifolius</i> (Ácido gálico)	800-1200		2,02	2,54	2,70	0,47	1,43
<i>A. colubrina</i> (catequina)	400-1200		0,22	-0,22	6,89	0,34	1,05

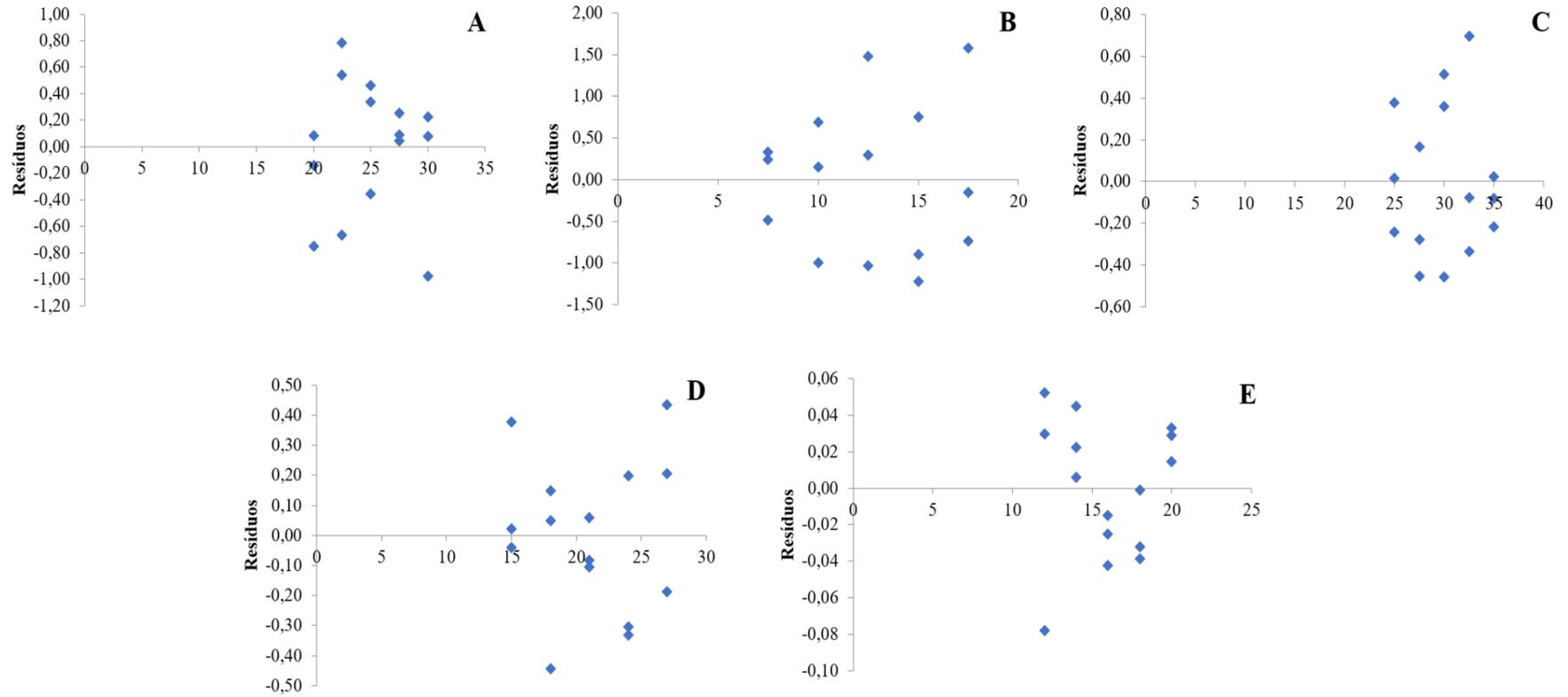
Onde: LD: Limite de detecção; LQ: Limite quantificação, em µg/mL.

Fonte: Autoria própria.

De acordo com análise estatística, verificou-se que os valores das áreas variaram em função da concentração havendo homocedasticidade nas variações. Tal veracidade pode ser confirmada ao analisar os gráficos de resíduos dos padrões polifenólicos (Figura 59) e dos marcadores químicos das espécies estudadas (Figura 60), onde os pontos foram distribuídos de forma equilibrada acima e abaixo da linha paralela ao eixo x na altura do resíduo zero formando uma faixa aproximadamente retangular.

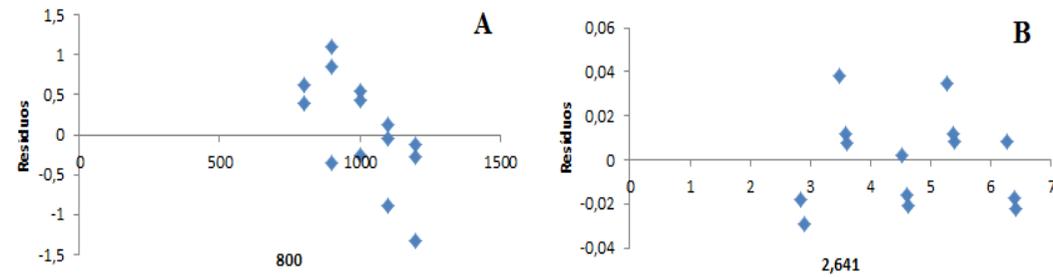
Dessa forma, as exigências preconizadas pela RDC 166/2017 para o parâmetro de proporcionalidade quando duas variáveis são correlacionadas foram estabelecidas (BRASIL, 2017). Os valores de R^2 para todas as curvas de calibração foram superiores a 0,99 e os valores de desvio padrão foram inferiores a 5%. As figuras 61 e 62 mostram as curvas de calibração e os valores de R^2 obtidos para os padrões de polifenóis e amostras de *S. terebinthifolius* e *A. colubrina*, respectivamente.

Figura 59 – Gráficos de resíduos obtidos para as curvas de calibração dos padrões de polifenóis estudados.



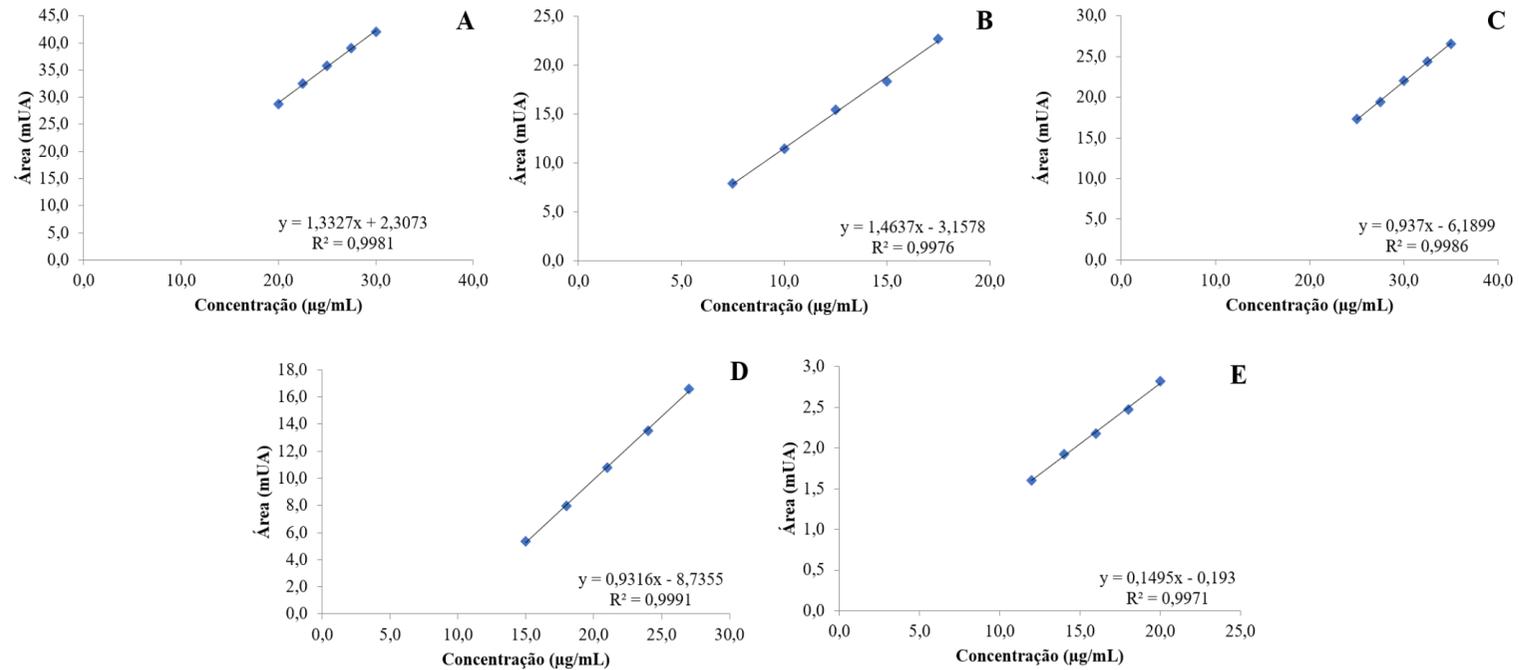
Onde: A – Ácido gálico; B – Quercetina; C – Ácido caféico; D – Rutina; E – Catequina.
 Fonte: Autoria própria.

Figura 60 - Gráficos de resíduos obtidos para as curvas de calibração de ácido gálico em *S. terebithinifolia* (A) e catequina em *A. colubrina* (B).



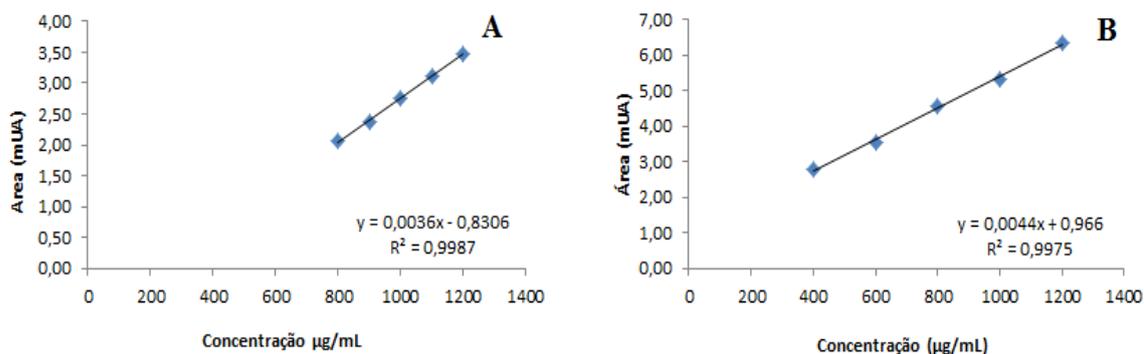
Fonte: Autoria própria.

Figura 61 - Curvas de calibração obtidas por CLAE dos padrões de polifenóis estudados.



Onde: A – Ácido gálico; B – Quercetina; C – Ácido caféico; D – Rutina; E – Catequina. Fonte: Autoria própria.

Figura 62 - Curvas de calibração obtidas por CLAE de ácido gálico em *S. terebinthifolia* (A) e catequina em *A. colubrina* (B).

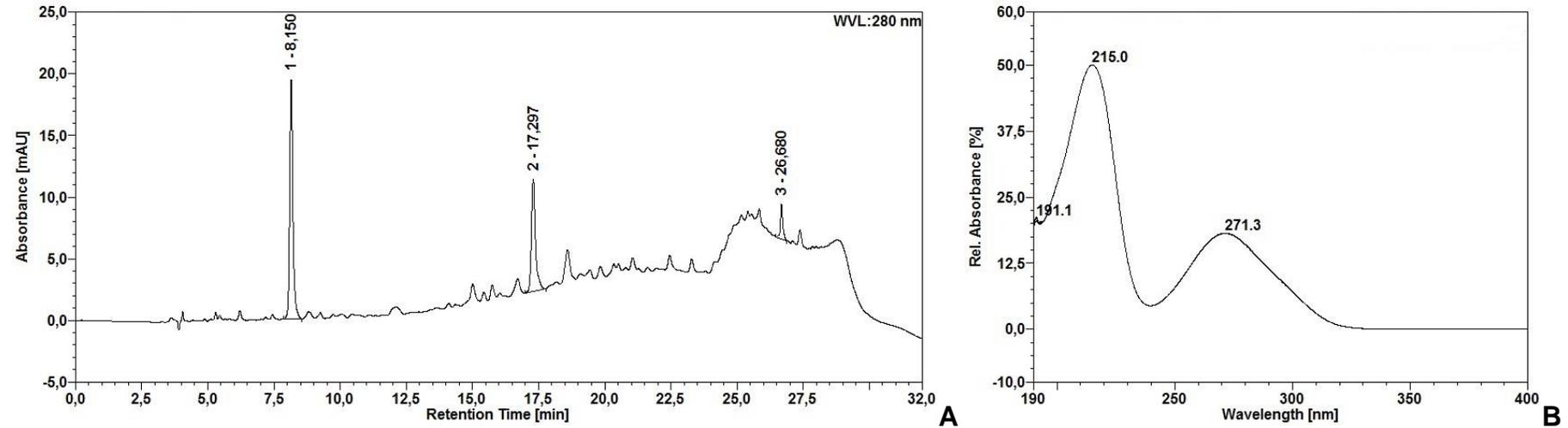


Fonte: Autoria própria.

Adicionalmente, os *fingerprints* obtidos por CLAE para as espécies estudadas estão apresentados abaixo. Como pode ser visualizado, os picos apresentaram simetria para os cromatogramas da fração acetato de etila de *S. terebinthifolius* (Figura 63A) e extrato bruto de *A. colubrina* (Figura 64A) permitindo reconhecer e identificar os marcadores químicos de interesse em ambas as espécies.

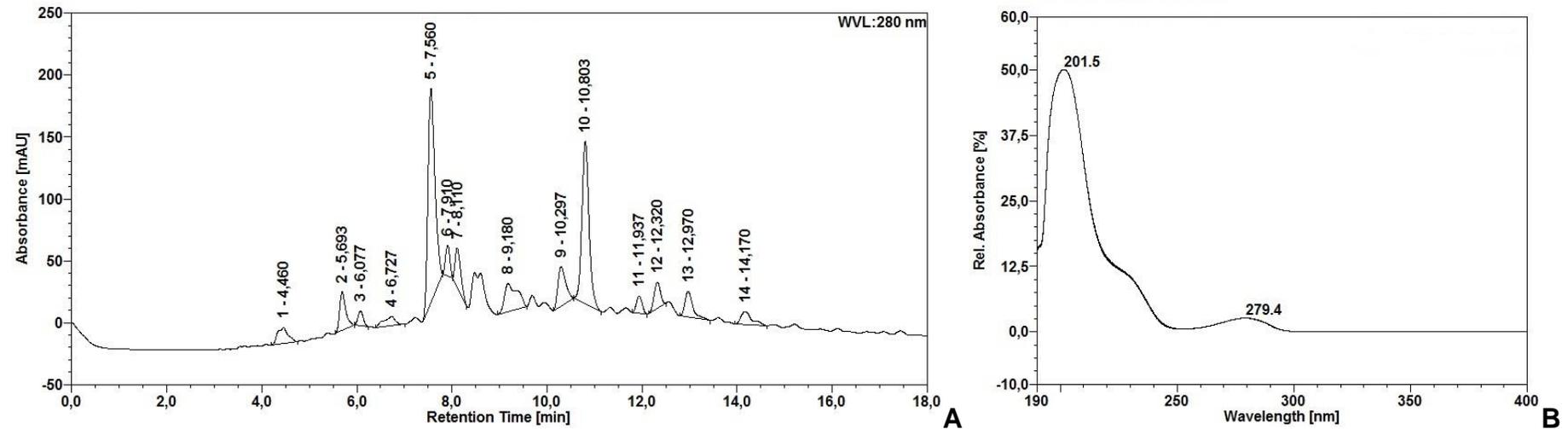
A identificação dos compostos bioativos de interesse foi realizada por comparação com os tempos de retenção e respectivos espectros de varredura. As análises cromatográficas da fração acetato de etila de aroeira permitiram identificar o ácido gálico, com tempo de retenção de aproximadamente 8,15 minutos (Figura 63B); e catequina para o extrato bruto de angico com tempo de retenção aproximadamente de 5,69 minutos (Figura 64B). A confirmação da identidade foi realizada pela comparação com os dados cromatográficos dos padrões.

Figura 63 – Cromatograma obtido por CLAE da fração acetato de etila de *S. terebinthifolius* em 280 nm.



Fonte: Autoria própria.

Figura 64 – Cromatograma obtido por CLAE do extrato bruto de *A. colubrina* em 280 nm.



Fonte: Autoria própria.

5 CONCLUSÃO

Diante dos resultados obtidos até o momento, pôde-se concluir acerca da viabilidade da utilização da ferramenta de imagem associada à rede neural artificial para quantificação de fitoconstituintes em espécies vegetais. Os testes realizados (divisão percentual e validação cruzada) e os modelos de Kernel empregados possibilitaram a correlação dos dados matemáticos fornecidos pela análise das imagens com os teores dos compostos estabelecidos por CLAE-DAD.

Para a maioria dos padrões de compostos polifenólicos foi observado que os modelos sigmoide, linear e polinomial apresentaram, respectivamente, os menores valores de erro percentual médio. Portanto, os modelos foram mais eficientes para descrever a correlação entre os dados oriundos do processamento das imagens, em ambos os comprimentos de onda (254 e 366 nm). Ademais, a elaboração dos gráficos de Boxplots possibilitou a interpretação estatística dos resultados de modelagem através da determinação dos desempenhos e dos menores valores de desvio padrão.

Por fim, os resultados alcançados até o momento foram promissores e sugerem que a abordagem analítica estudada atende ao objetivo de quantificação de compostos do metabolismo secundário empregando processamento de imagem. Entretanto, a avaliação dos compostos nas matrizes vegetais ainda requer outras adequações operacionais e matemáticas, para que ferramenta possa ser adotada na rotina do controle de qualidade de matérias-primas e produtos acabado.

REFERÊNCIAS

ABU-REIDAH, I.M.; ALI-SHTAYEH, M.S.; JAMOUS, R.M.; ARRÁEZ-ROMÁN, D.; SEGURA-CARRETERO, A. HPLC–DAD–ESI-MS/MS screening of bioactive components from *Rhus coriaria* L. (Sumac) fruits. **Food Chemistry**, v. 166, p. 179–191, 2015.

AZEVEDO, W. W.; LIMA, S. M. L.; FERNANDES, I. M. M.; ROCHA, A. D. D.; CORDEIRO F. R.; SILVA-FILHO, A. G.; SANTOS, W. P. Morphological Extreme Learning Machines Applied to Detect and Classify Masses in Mammograms. **International Joint Conference on Neural Networks**, p. 1-8, 2015).

ABOU-DONIA, A. H.; DARWISH, F. A.; TOAIMA, S. M.; SHAWKY, E.; TAKLA, S. S. A new approach to develop a standardized method for assessment of acetylcholinesterase inhibitory activity of different extracts using HPTLC and image analysis. **Journal of Chromatography B**, v. 955, p. 50-57, 2014.

ALMEIDA A.C.; BARROS, P. L. C.; MONTEIRO, J.H.A.; ROCHA, B.R.P. Estimation of above ground forest biomass in Amazonia with neural networks and remote sensing. **IEEE Latin America Transactions**, v. 7, n. 1, p. 27-32, 2009.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: LTC, 2 ed. 2007.

BEZERRA, I. C. F. **Investigação Fitoquímica de Extrato Bruto e Frações das Folhas de *Eugenia uniflora* L. (Pitangueira)**. 2016. Dissertação de mestrado, Universidade Federal de Pernambuco, 2016. 132 p.

BINOTI, D.H.B.; BINOTI, M. L. M. S.; LEITE, H.G. Configuração de redes neurais artificiais para estimação do volume de árvores. **Brazilian Journal. Wood Science**, v. 5, n. 1, p. 58-67.

BOSCOLO, O. H.; VALLE, L. S. Plantas de uso medicinal em Quissamã, Rio de Janeiro, Brasil. **Série Botânica**, Porto Alegre, v. 63, n. 2, p. 263-277, 2008.

BURQUE, R.K.; FRANCESCONI, L.P.; VICTORINO, A.T.; MASCARENHAS, A.A.; CERESÉR, K.M. Determinação de compostos fenólicos e avaliação da atividade antioxidante de *Lafoensia pacari* (LYTHRACEAE). **Revista eletrônica de farmácia**, v. 12, n. 1, p. 1-10, 2015.

CAMBRIA, E.; HUANG, G. B.; KASUN, L.L.C.; ZHOU, H.; C.; VONG, C. M.; YING, J.; et. al., J. Extreme learning machines, **IEEE Intelligent Systems**, v. 28, n. 6, p. 30–59, 2013.

CUSTÓDIO, A.T.N. **Análise de imagens de imunohistoquímica da placenta e do pulmão usando segmentação e descritores de forma**. Trabalho de Conclusão de Curso Engenharia da Computação. Escola Politécnica de Pernambuco. UPE. Orientador: Prof. Wellington Pinheiro dos Santos. 2009

DECONINCK, E.; SACRE, P. Y.; COURSELLE, P.; DE BEER, J. O. Chromatography in the detection and characterization of illegal pharmaceutical preparations. **Journal of Chromatography Science**, v. 51, p. 791–806, 2013.

DOUGLAS, C.; PECK, A.; G. GEOFFREY V. G. Introduction to Linear Regression Analysis, 5th Edition. Editora Wiley, 2012.

ERSAN, S.; GUCLU USTUNDAG, O.; CARLE, R.; SCHWEIGGERT, R.M. Identification of phenolic compounds in red and green pistachio (*Pistacia vera* L.) hulls (Exo- and Mesocarp) by HPLC-DAD-ESI-(HR)-M(n)S. **Journal of Agricultural and Food Chemistry**, v. 64, p. 5334–5344, 2016.

EVERETTE, J. D.; BRYANT, Q. M.; GREEN, A. M.; ABBEY, Y. A.; WANGILA, G. W.; WALKER, R. B. Thorough study of reactivity of various compound classes toward the Folin-Ciocalteu reagent. **Journal of Agricultural and Food Chemistry**, v. 58, p. 8.139-8.144, 2010.

FERREIRA, M.R.; FERNANDES M.T.; SILVA, W.A.; BEZERRA I.C.; SOUZA, T.P.; PIMENTEL, M.F.; SOARES, L.A.L. Chromatographic and Spectrophotometric

Analysis of Phenolic Compounds from Fruits of *Libidibia ferrea* Martius. **Pharmacognosy Magazine**, v. 12, p. 285-291, 2016.

FLEMING, G. M. Design of a high resolution image cytometer with open software architecture. **Analytical Cellular Pathology**, v. 10, n. 1, p. 1-11, 1996. Disponível em <
http://murphylab.web.cmu.edu/publications/boland/boland_node26.html#fig:haralick_neighbors> Acesso em: maio, 2019.

GUERRA, A. M. N. M.; PESSOA, M. F.; SOUZA, C. S. M.; MARACAJÁ, P. B. Utilização de plantas medicinais pela comunidade rural Moacir Lucena, Apodi – RN. **Journal of Biosciences**, v. 26, n. 3, p. 442-450, 2010.

GULZEMEREC, E.; RISTIVOJEVIC, P.; VOVK, I.; MILOJKOVIC-OPSENICA, D.; YESILADA, E. Quality assessment of marketed chamomile tea products by a validated HPTLC method combined with multivariate analysis. **Journal of Pharmaceutical and Biomedical Analysis**, v. 132, p. 35-45, 2017.

GLÓRIA, L.L., ARANTES, M.B.S., PEREIRA, S.M.F., VIEIRA, G.S., MARTINS, C.X., JUNIOR, A.R.C., ANTUNES, F., BRAZ-FILHO, R., VIEIRA, I.J.C., CRUZ, L.L., CHAVES, D.S.A., FREITAS, S.P., BARROS DE OLIVEIRA, D.B. Phenolic compounds present *Schinus terebinthifolius* Raddi influence the lowering of blood pressure in rats. **Molecules**, v. 22, 2017.

HASLAM, E. Vegetable tannins—Lessons of a phytochemical lifetime. **Phytochemistry**, v. 68, n. 22, p. 2713-2721, 2007.

HAWRYŁ, A., et al. TLC Profiles of Selected Cirsium Species with Chemometrics in Construction of Their Fingerprints. **Journal of Chromatographic Science**, v. 54, n. 7, p. 1096-104, 2016.

HUANG, G. B.; WANG, D. H.; LAN, Y. Extreme learning machines: a survey, **International Journal of Machine Learning and Cybernetics**, v. 2, n. 2, p. 107–122, 2011.

HUANG, G. B.; ZHOU, H.; DING, X.; ZHANG, R. Extreme learning machine for regression and multiclass classification. **IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics**, v. 42, n. 2, p. 513–529, 2012.

JAFFEE, E. M.; DANG, C. V.; AGUS, D. B.; ALEXANDER, B. M.; ANDERSON, K. C.; ASHWORTH, A. Future cancer research priorities in the USA: a lancet oncology commission. **The Lancet Oncology**, v. 18, n. 11, p. 653–706. 2017.

JOAQUIM, C. F.; OLIOTI, J.B.; TRAINA, A.J.M. Discriminação de Aspectos Malignos em Massas Tumorais de Mamografias Usando Características de Forma das Imagens. IV SBQS - V Workshop de Informática Médica (2005).

JOHN W. EATON, DAVID BATEMAN, SOREN HAUBERG, RIK WEHBRING (2015). GNU Octave version 4.0.0 manual: a high-level interactive language for numerical computations. URL <http://www.gnu.org/software/octave/doc/interpreter/>.

LORENZI, H.; MATOS, F. J. A. **Plantas medicinais no Brasil: nativas e exóticas**. 2. Ed. São Paulo: Instituto Plantarum, 2008. 560 p.

KEATING, E.; MARTEL, F. Antimetabolites effects of polyphenols in breast cancer cells: focus on glucose uptake and metabolism. **Polyphenols and breast cancer**. v. 5, p.1-11, 2018.

KIM, D.H.; KHAN, H.; ULLA H, H.; ŠMEJKAL, K.; EFFERTH, T.; MAHAMOODALLY, M.F.; XU, S.; HABERMARIAM, S.; FILOSA, R.; LAGOA, R.; RENGASAMY, K.R. MicroRNA visando a **quercetina** no tratamento de câncer e quimioproteção. **Pharmacology Research**, 2019.

KOSURU, R.Y.; ROY, A.; DAS, S.K.; BERA, S. Gallic acid and gallates in human health and disease: Do mitochondria hold the key to success? **Molecular Nutrition Food Research**, v. 62, n. 1, 2017.

KUMAR, S., CHANDRA, P., BAJPAI, V., SINGH, A., SRIVASTAVA, M., MISHRA, D.K., KUMAR, B. Rapid qualitative and quantitative analysis of bioactive compounds from *Phyllanthus amarus* using LC/MS/MS techniques. **Industrial Crops and Products**, v. 69, p. 143–152, 2015.

LEITE, A.S.; ISLAM, T.M.D.; JUNIOR, A.L.G.; SOUSA, J.M.C.; ALENCAR, M.V.O.; PAZ, M.F.C.J. et al. Pharmacological properties of cashew (*Anacardium occidentale*). **African Journal of Biotechnology**, v. 13, n.35, p. 1855-1863, 2016.

LIAO, S.G.; ZHANG, L.J.; SUN, F.; WANG, Z.; HE, X.; WANG, A.M.; LI, Y.J.; HUANG, Y.; LAN, Y.Y.; ZHANG, B.L.; WANG, Y.L. Identification and characterisation of phenolics in *Polygonum capitatum* by ultrahigh-performance liquid chromatography with photodiode array detection and tandem mass spectrometry. **Phytochemical Analysis**, v. 24, p. 556–568. 2013.

LIMA, R.F.; ALVES, E.P.; ROSALEN, P.L.; RUIZ, A.L.T.G.; DUARTE, M.C.T.; GÓES, V.F.F.; MEDEIROS, A.C.D.; PEREIRA, J.V.; GODOY, G.P.; COSTA, E.M.M.B. Antimicrobial and Antiproliferative Potential of *Anadenanthera colubrina* (Vell.) Brenan. **Evidence-Based Complementary and Alternative Medicine**, v. 2014, p.1-7, 2014.

LOESCHER, C.M.; MORTANA, D. W.; RAZIC, S.; KUSTRIN, S. A. High performance thin layer chromatography (HPTLC) and high performance liquid chromatography (HPLC) for the qualitative and quantitative analysis of *Calendula officinalis*—Advantages and limitations. **Journal of Pharmaceutical and Biomedical Analysis**, v. 98, p. 52-59, 2014.

LORENZI, H. **Árvores Brasileiras: Manual de identificação e cultivo de plantas arbóreas nativas do Brasil**. Brazilian Trees. Identification and cultivation Manual of native tree plants of Brazil. Instituto Plantarum; São Paulo; 2009.

LOSADA-ECHEBERRIA, M.; HERRANZ-LÓPEZ, M.; MICOL, V.; BARRAJÓN-CATALÁN, E. Polyphenols as promising drugs against main breast cancer signatures. **Antioxidants**, v. 6, n. 4, p. 1-24, 2017.

MARCH, R.E.; LEWARS, E.G.; STADEY, C.J.; MIAO, X.S.; ZHAO, X.; METCALFE, C.D. A comparison of flavonoid glycosides by electrospray tandem mass spectrometry. **International Journal Mass Spectrometry**, v. 248, p. 61–85, 2015.

MARCHESI, J.R.; ADAMS, D.H.; FAVA, F.; HERMES, G.D.A.; HIRSCHFIELD, G.M.; HOLD, G.; QURAIISHI, M.N.; KINROSS, J.; SMIDT, H.; TUOHY, K.M.; THOMAS, L.V.; ZOETENDAL, E.G.; HART, A. The gut microbiota and host health: a new clinical frontier. **Gut**, v. 65, n. 2, p. 330–9, 2016.

MARSTON, A. Thin layer chromatography with biological detection in phytochemistry. **Journal Chromatography A**, v. 1218, n. 19, p. 2676-2683, 2011.

MOKAC, J. Chemometrics in medicine and pharmacy. **Nova Biotechnologica et Chimica**, v. 11, n.1, p. 11-25, 2012.

MONDAL, S.; DEBJIT, G.; GANAPATYA, S.; CHEKUBOYINA, S. V. G.; SAMALA, M. Hepatoprotective activity of *Macrothely pteristorresiana* (Gaudich.) aerial parts against CCl₄-induced hepatotoxicity in rodents and analysis of polyphenolic compounds by HPTLC. **Journal of Pharmaceutical Analysis**, v. 7, p. 181-187, 2017.

MONTEIRO, J. M.; ALBUQUERQUE, U. P.; ARAÚJO, E. L. Taninos: uma abordagem da química à ecologia. **Química Nova**, v. 28, n. 5, p. 892-896, 2005.

GRACIENE, S. M.; CAROLINE, J. S.; ISABEL M.; QUILHO, T.; MORI, A. F.; PEREIRA, H. Bark anatomy, chemical composition and ethanol-water extract composition of *Anadenanthera peregrina* and *Anadenanthera colubrina*. **PLOS one**, v. 27, p.1-14, 2017.

MUCCILLI, V.; CARDULLO, N.; SPATAFORA, C.; CUNSOLO, V.; TRINGALI, C. α -Glucosidase inhibition and antioxidant activity of an oenological commercial tannin. Extraction, fractionation and analysis by HPLC/ESI-MS/MS and (1) HNMR. **Food Chemistry**, v. 215, p. 50–60, 2017.

OLECH, M.; KOMSTA, L.; NOWAK, R.; GIESLA, L.; MONICA, W.H. Investigation of antiradical activity of plant material by thin-layer chromatography with image processing. **Food Chemistry**, v. 132, p. 549-553, 2012.

OLIVEIRA, L.F.; NARLOCH, A.L.M.; KIST, D.M.; SOARES FILHO, M.P.; MENEGHELLO, G.E.; CAVALHEIRO, G.G.H.; TILMANN, M.A.A. Extração de Características de Forma utilizando matriz de co-ocorrência e Atributos de Haralick. Disponível em http://iris.sel.eesc.usp.br/wvc/Anais_WVC2012/pdf/97967.pdf Acesso em maio, 2019.

RAJAGOPAL, C.; LANKADASARI, M.B.; ARANJANI, J.M.; HARIKUMAR, K.B. Targeting oncogenic transcription factors by polyphenols: a novel approach for cancer therapy. **Pharmacological Research**, v. 130, p. 273-291, 2018.

RENISUS, MS - Ministério da Saúde. **Relação de Plantas Medicinais de Interesse ao SUS**. Brasília, 2009.

- ROCHA, W. S.; LOPES, M. L.; SILVA, D. B.; VIEIRA, R. F.; SILVA, J.P.; AGOSTINI-COSTA, T.S. Compostos fenólicos totais e taninos condensados em frutas nativas do cerrado **Revista Brasileira de Fruticultura**, v. 33, n. 4, p. 1215-1221, 2011.
- ROCHA, P.S.; BOLETI, A. P. A.; VIEIRA, M. C.; CAROLLO, C. A.; SILVA, D. B. S.; ESTEVINHO, L. M.; SANTOSA, E. L; SOUZA, K.P. Microbiological quality, chemical profile as well as antioxidant and antidiabetic activities of *Schinus terebinthifolius* Raddi. **Comparative Biochemistry and Physiology**, parte C v. 220, p. 36-46, 2019.
- ROSASCO, L. Reproducing Kernel Hilbert Spaces - MIT 9.520 Statistical Learning Theory and Applications, Lecture Notes. 10 fev. 2010. Disponível em: <<http://www.mit.edu/~9.520/fall14/Classes/rkhs.html>>. Acesso em: 16 maio 2018.
- SANTANA, M.A.; PEREIRA, J.M.S.; SILVA, F.L.; LIMA, N.M.; SOUSA, F.N.; ARRUDA, G.M.S.; LIMA, R.C.F.; SILVA, W.W.A.; SANTOS, W.P. Breast cancer diagnosis based on mammary thermography and extreme learning machines. **Research on biomedical engineering**, v. 34, p. 45-53, 2018.
- SANTOS, W. P. **Análise Digital de Imagens em Patologia: Ferramentas de Morfologia Matemática e Lógica Fuzzy**. New York: Create Space, 2009.
- SANTOS, J.S.; MARINHO, R.R.; EKUNDI-VALENTIM ET AL. E, Beneficial effects of *Anadenanthera colubrina* (Vell.) Brenan extract on the inflammatory and nociceptive responses in rodent models. **Journal of Ethnopharmacology**, v.148, p. 218–222, 2013.
- SARTORI, C.J.; MORI, F.A.; VALLE, M.L.A.; MENDES, L.M.; PROTASIO, T.P. Gravimetric yield in condensed tannins in the barks of *Anadenanthera peregrina* in different diametric classes. **Cerne**, v. 20, p. 239–244, 2014.
- SCHNEIDER, C.A.; RASBAND, W. S.; ELICEIRI, K. W. NIH Image to imageJ: 25 years of image analysis. **Nature methods**, v. 9, n.7, p. 671-675, 2012.
- SENGER, A. E. V.; SCHWANKE, C. H. A.; GOTTLIEB, M. G. V. Chá verde (*Camellia sinensis*) e suas propriedades funcionais nas doenças crônicas não transmissíveis. **Scientia Medica**, v. 20, n. 4, p. 292-300, 2010.
- SHERMA, J. Biennial review of planar chromatograph: 2009-2011. **Journal of AOAC international**, v. 95, n. 4, p. 992-1009, 2012.
- SHEWIYO, D.H. HPTLC methods to assay active ingredients in pharmaceutical formulations: A review of the method development and validation steps. **Journal of Pharmaceutical and Biomedical Analysis**, v. 66, p. 11–23, 2012.
- SILVA, J.H.S.; SIMAS, N.K.; ALYANO, C.S.; ALVIANO, D.S.; VENTURA, J.A.; LIMA, E.J., SEABRA, S.H.; KUSTER, R.M. Anti- *Escherichia coli* activity of extracts from *Schinus terebinthifolius* fruits and leaves. **Natural Product Research**, v. 32, n. 2, p.1-4, 2017a.
- SILVA, M.M.; IRIGUCHI, E.K.K.; KASSUYA, C.A.L.; VIEIRA, M.C.; FOGLIO, M.A.; CARVALHO, J.E.; RUIZ, A.L.T.G.; SOUZA, K.P.; FORMAGIO, A.S.N. *Schinus terebinthifolius*: phenolic constituents and in vitro antioxidant, antiproliferative and in

vivo anti-inflammatory activities. **Revista Brasileira de Farmacognosia**, v. 27, p. 445–452, 2017b.

SIMA, I. A.; CASONI, D.; SÂRBU, C. High sensitive and selective HPTLC method assisted by digital image processing for simultaneous determination of catecholamines and related drugs. **Talanta**, v. 114, p. 117-123, 2013.

SIMÕES, C. M. O. et al. **Farmacognosia do Produto Natural ao Medicamento**. 1. ed. 502p. Artmed, 2016.

SIMÕES, V.N.; FAVARIN, L. R. V.; CABEZA, N. A. OLIVEIRA, T. D.; FIORUCCI, A. R.; STROPA, J. M.; RODRIGUES, D. C. M.; CAVALHEIRO, A. A.; ANJOS, A. Síntese, caracterização e estudo das propriedades de um novo complexo mononuclear contendo quercetina e íon ga(iii). **Química Nova**, v. 36, n.4, p. 495-501, 2013.

SINGH, I.; JUNEJA, P.; KAUR, B.; KUMAR, P. Pharmaceutical Applications of Chemometric Techniques. **ISRN Analytical Chemistry**, v. 2013, p. 1-13, 2013.

SOARES, L. A. L.; FERREIRA, M. R. A. **Standardization and Quality Control of Herbal Medicines**. In.: FREITAS, L.A.P.; TEIXIERA, C.C.C.; ZAMARIOLI, C.M. et al. Recent Development in Phytomedicine Technology. New York: Nova Science Publishers Inc., Chap. 9, 2017.

SOUZA, L. F. L. **Atividade antimicrobiana de extratos de Aroeira (*Schinus terebinthifolius* Raddi) frente a bactérias relacionadas à mastite bovina**. Brasília: Distrito Federal, 2011. Disponível em < <https://docplayer.com.br/18622967-Atividade-antimicrobiana-de-extratos-de-aroeira-schinus-terebinthifolius-raddi-frente-a-bacterias-relacionadas-a-mastite-bovina.html> > Acesso em: 15 de jul. de 2019.

SUN, J.; CHEN, P. Chromatographic fingerprint analysis of yohimbe bark and related dietary supplements using UHPLC/UV/MS. **Journal of Pharmaceutical Biomedical Analysis**, v. 6, p. 142–149, 2012.

TISTAERT, C.; DEJAEGHER, B.; HEYDEN, Y. V. Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. **Analytica Chimica Acta**, v. 690, p. 148-161, 2011.

WILLIAMSON, G. The role of polyphenols in modern nutrition. **Nutrition Bulletin**. v. 42, n. 3, p. 226–235, 2017.

WONG, K. H. et al. Differentiating *Puerariae Lobatae* Radix and *Puerariae Thomsonii* Radix using HPTLC coupled with multivariate classification analyses. **Journal of Pharmaceutical and Biomedical Analysis**, v. 95, p. 11–19, 2014.

WU, Y.H.; ZHANG, B.Y.; QIU, L.P.; GUAN, R.F.; YE, Z.H.; YU, X.P. Structure Properties and Mechanisms of Action of Naturally Originated Phenolic Acids and Their Derivatives against Human Viral Infections. **Current Medicinal Chemistry**, v. 24, p. 4279 – 4302, 2017.

XIAO, J.B. Polyphenol-Plasma Proteins Interaction: Its Nature, Analytical Techniques, and Influence on Bioactivities of Polyphenols. **Current. Drug Metabolism**, v. 14, p. 367–368, 2013.

XIE, Y.; CHEN, J.; XIAO, A.; LIU, L. Antibacterial Activity of Polyphenols: Structure-Activity Relationship and Influence of Hyperglycemic Condition. **Molecules**, v. 22, p. 1913-1924, 2017.

YANG, C.S.; LAMBERT, J. D.; SANG, S. Antioxidative and anti-carcinogenic activities of tea polyphenols. **Archives of Toxicology**, v. 83, p. 11-21, 2010.

ZANIN, J. L. B., et al. The Genus *Caesalpinia* L. (Caesalpinaceae): Phytochemical and Pharmacological Characteristics. **Molecules**, v. 17, p. 7887-7902, 2012.

ZHANG, T.; ZHONG, S.; HOU, L.; LI, T.; XING, X. J.; GUAN, T.; ZHANG, J.; WANG, Y. Estrogenic properties of coumarins and meroterpene from the fruits of *Cullen corylifolium*: Experimental and computational studies. **Phytochemistry**, v. 152, p. 148-153, 2018.

ZHU, J.; JIANG, J. Pharmacological and Nutritional Effects of Natural Coumarins and Their Structure–Activity Relationships. **Molecular nutrition food research**, p. 1-11, 2018.

ANEXO A - Comprovante depósito de patente

Data: seg, 22 de jul de 2019 às 17:04
Assunto: SIPAC - Informativo Movimentação de Processo 23076.034623/2019-16
Para: SIPAC+-+PROCESSOS <rhayanne.moraes11@gmail.com>

Prezado(a) RHAYANNE THAIS DE MORAES RAMOS,

Uma nova movimentação foi efetuada para o processo 23076.034623/2019-16 do qual o sr(a). faz parte como interessado.

Dados Gerais do Processo:

Processo: 23076.034623/2019-16

Tipo do Processo: REGISTRO DE PATENTE

Assunto Detalhado: REGISTRO DE PATENTE DE DESENVOLVIMENTO DE SISTEMA DE ANÁLISE DE IMAGENS PARA QUANTIFICAÇÃO DE COMPOSTOS POLIFENÓLICOS A PARTIR DE FINGERPRINTS POR CROMATOGRAFIA EM CAMADA DELGADA

Natureza do Processo: RESTRITO

Dados Gerais da Movimentação:

Tipo de Movimentação: RECEBIMENTO

Data de Recebimento: 22/07/2019 17:04

Unidade de Destino: POSITIVA - DIRETORIA DE INOVACAO - GR (11.01.18)

Unidade de Origem: COORDENACAO DA POS-GRADUACAO EM INOVACAO TERAPEUTICA - CB (11.84.05)

##MENSAGEM_INFORMACAO_CONSULTA_PROCESSO##

Esta mensagem foi gerada automaticamente pelo sistema e não deve ser respondida.

—
Rhayanne Moraes
Farmacêutica
Mestranda em Inovação Terapêutica - UFPE
(81) 9.8467-5982

