

UNIVERSIDADE NOVE DE JULHO - UNINOVE
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

JORGE HENRIQUE PESSOTA

**SISTEMA ESPECIALISTA APLICADO À INSPEÇÃO DA QUALIDADE VISUAL DE
GRÃOS DE FEIJÃO**

SÃO PAULO

2013

JORGE HENRIQUE PESSOTA

**SISTEMA ESPECIALISTA APLICADO À INSPEÇÃO DA QUALIDADE VISUAL DE
GRÃOS DE FEIJÃO**

Dissertação apresentada ao Programa de Pós Graduação em Engenharia de Produção da Universidade Nove de Julho – UNINOVE, como exigência parcial para obtenção do título de Mestre em Engenharia de Produção.

Orientador: Prof. Dr. Sidnei Alves de Araújo

SÃO PAULO

2013

Pessota, Jorge Henrique.

Sistema especialista aplicado à inspeção da qualidade visual de grãos de feijão. /Jorge Henrique Pessota. 2013.

83 f.

Dissertação (mestrado) – Universidade Nove de Julho - UNINOVE, São Paulo, 2013.

Orientador (a): Prof. Dr. Sidnei Alves de Araújo.

1. Sistema especialista. 2. Inspeção automática. 3. Qualidade visual de alimentos

I. Araújo, Sidnei Alves. II. Título

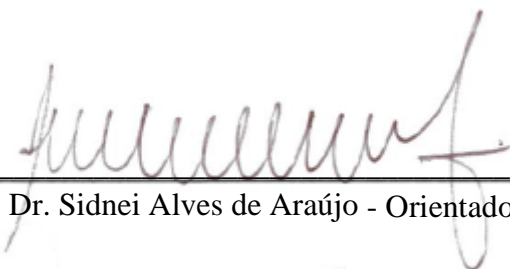
CDU 658.5

**SISTEMA ESPECIALISTA APLICADO À INSPEÇÃO DA QUALIDADE VISUAL DE
GRÃOS DE FEIJÃO**

Por

JORGE HENRIQUE PESSOTA

Dissertação apresentada ao Programa de Pós Graduação em Engenharia de Produção da Universidade Nove de Julho – UNINOVE, como exigência parcial para obtenção do título de Mestre em Engenharia de Produção.



Presidente: Prof. Dr. Sidnei Alves de Araújo - Orientador, UNINOVE



Membro interno: Prof. Dr. André Felipe Henriques Librantz, UNINOVE



Membro externo: Prof. Dr. Rodrigo Franco Gonçalves, POLI-USP/UNIP

SÃO PAULO

2013

“Algo só é impossível até que alguém duvide e acabe provando o contrário.”
(Albert Einstein)

AGRADECIMENTOS

Agradeço a Deus pela força e saúde concedida para superar os desafios que surgem a cada dia. À minha esposa e família pela paciência e compreensão durante os períodos de ausência. Ao meu orientador Prof. Dr. Sidnei Alves de Araújo pelo apoio e companheirismo empregado ao longo desta jornada. Ao Prof. Ms. Wonder Alexandre Luz Alves pelo incentivo em iniciar este desafio e à UNINOVE pela oportunidade oferecida.

RESUMO

As propriedades visuais dos alimentos são importantes fatores para a definição da sua qualidade e para a determinação do seu preço de mercado. Por esse motivo, o processo de inspeção da qualidade visual dos alimentos precisa ser confiável, de modo que atenda às expectativas dos consumidores e às normas estabelecidas pelo governo. No entanto, é comum que esta inspeção ocorra de maneira manual, a qual pode demandar muito tempo, ser tediosa, gerar altos custos operacionais, ser passível a falhas humanas e apresentar dificuldades para padronizar os resultados, tornando desejável a sua automatização. No caso do feijão comercializado no Brasil, sua qualidade está relacionada com a quantidade de defeitos e de misturas presentes nas amostras avaliadas e sua classificação é feita manualmente, com base em normas e procedimentos estabelecidos pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Neste trabalho foi desenvolvido um Sistema Especialista (SE) aplicado à inspeção da qualidade visual de grãos de feijão, com o objetivo de classificá-los de acordo com a coloração de suas películas. Para tanto, técnicas de reconhecimento de padrões visuais e de inteligência computacional foram empregadas. Também foi proposto um método para detecção dos grãos visando superar uma limitação presente nos sistemas de reconhecimento de padrões visuais disponíveis na literatura, voltados para a análise e classificação de sementes e grãos. Esta limitação faz com que os grãos sejam distanciados uns dos outros para facilitar o processo de segmentação e detecção, o que dificulta a sua aplicabilidade em processos industriais. Nos experimentos realizados foi empregado um aparato composto por uma esteira automática com uma câmera acoplada, com o intuito de tornar o ambiente experimental mais próximo ao encontrado em um processo produtivo. As taxas médias de acertos (98,62%), precisão (98,50%) e acurácia (97,16%) obtidas nos experimentos indicam que o SE desenvolvido pode ser incorporado ao processo industrial de inspeção da qualidade visual de grãos de feijão.

Palavras-chave: sistema especialista, inspeção automática, qualidade visual de alimentos, feijão, reconhecimento de padrões, inteligência computacional.

ABSTRACT

The visual properties of foods are an important factor for the definition of their quality and for determining their market price. For this reason, the inspection process for visual quality of food must be reliable, in order to attend to consumer expectations and to the standards set by the government. However, it is usual that this kind of inspection occurs manually and, if it happens, it can be slow and tedious; generate high operating costs; be subject to human failings, becoming the results standardization more difficult and making its automation desirable. In the case of the beans traded in Brazil, their quality is related to the amount of defects and mixtures that they can present in the evaluated samples. Their classification is made manually; it is based on rules and procedures established by the Ministry of Agriculture, Livestock and Food Supply (Ministério da Agricultura, Pecuária e Abastecimento - MAPA). In this research work, it was developed an Expert System (Sistema Especialista - SE) applied to the visual quality inspection of grains of bean, in order to classify them according to the color of their skin/coverage. For this purpose, techniques of visual pattern recognition and artificial intelligence were used. Additionally, it was proposed a method for grain detection with the purpose of overcoming a limitation that is present in the recognition system of visual patterns available in the literature, destined to the analysis and classification of seeds and grains. This limitation keeps the grains spaced from each other, in or order to facilitate the process of segmentation and detection, which becomes its applicability in industrial processes more difficult. In the experiments done it was employed an apparatus composed of an automatic conveyor belt with a camera attached, in order to make the experimental environment as close as possible to those found in a production process. The average rates of success (98.62%), precision (98.50%) and accuracy (97.16%) obtained in the experiments indicate that the developed Expert System can be incorporated into industrial process of visual quality inspection of grains of bean.

Keywords: expert system, automatic inspection, visual quality of food, beans, pattern recognition, computational intelligence.

LISTA DE FIGURAS

Figura 1: Componentes de um sistema de visão computacional.	24
Figura 2: Etapas do processamento de imagens digitais.	26
Figura 3: Exemplos de núcleos utilizados pelo método proposto por Maruta (2011).....	28
Figura 4: Redes alimentadas diretamente com camada única	31
Figura 5: Redes alimentadas diretamente com múltiplas camadas	31
Figura 6: Redes recorrentes	32
Figura 7: Feijões considerados neste trabalho.	36
Figura 8: Desenho esquemático do aparato.	36
Figura 9: Fluxograma dos experimentos realizados.	39
Figura 10: Aparato para aquisição das imagens e a interface gráfica para utilização do SE. ..	42
Figura 11: Cinco imagens com variação na iluminação, capturadas de uma mesma amostra.	44
Figura 12: Imagens da amostra após o cálculo da magnitude do gradiente.	45
Figura 13: Imagem Z.	46
Figura 14: Imagem S.	47
Figura 15: Núcleos gerados para aplicação da granulometria.	48
Figura 16: Resultado da aplicação da granulometria.....	48
Figura 17: Disposição dos grãos de feijão em uma amostra analisada. À direita a imagem original da amostra e à esquerda a imagem após sofrer o processo de segmentação.....	49
Figura 18: Disposição dos grãos de arroz em uma amostra analisada. À direita a imagem original da amostra e à esquerda a imagem após sofrer o processo de segmentação.....	49
Figura 19: Círculos concêntricos gerados para a extração dos atributos.....	50
Figura 20: Processo de classificação pelas RNA.....	51
Figura 21: Processo de classificação pela técnica K-MEANS.....	52
Figura 22: Classificação realizada pela RNA.....	53
Figura 23: Classificação realizada por K-MEANS.	53
Figura 24: Diferenças geradas na classificação a partir das técnicas RNA e K-MEANS, para uma mesma amostra.	54

Figura 25: Imagem com iluminação uniforme pertencente à amostra 5 do subconjunto 9.....	56
Figura 26: Imagem com iluminação uniforme pertencente à amostra 01 do subconjunto 3....	59
Figura 27: Taxa de acurácia para as técnicas de classificação RNA e K-MEANS.....	61
Figura 28: Comparativo de desempenho do SE utilizando RNA+RGB com e sem o esquema de iluminação e pré-processamento proposto.....	62

LISTA DE TABELAS

Tabela 1: Percentual de mistura tolerada em cada classe de feijão	21
Tabela 2: Subconjuntos de amostras gerados para a realização dos experimentos.	37
Tabela 3: Valores assumidos para os parâmetros durante os experimentos.....	39
Tabela 4: Resultados experimentais obtidos com classificação K-MEANS e sistemas de cores RGB.....	55
Tabela 5: Resultados experimentais obtidos com classificação K-MEANS e sistemas de cores CIELAB.....	57
Tabela 6: Resultados experimentais obtidos com classificação através da RNA e sistemas de cores RGB.	58
Tabela 7: Resultados experimentais obtidos com classificação através da RNA e sistemas de cores CIELAB.	60
Tabela 8: Resultados experimentais sumarizados.	61

LISTA DE ABREVIATURAS E SIGLAS

BC	Base de Conhecimento
C/C++	Linguagens de programação. C é uma linguagem estruturada, enquanto C++ é orientada a objetos.
CIE	<i>Commission Internationale 'd'Eclairage</i> .
CIELAB	Sistema de cores perceptualmente uniforme desenvolvido pela CIE. Também conhecido como $L^*a^*b^*$.
Dev-C++	Ambiente integrado de desenvolvimento de aplicações escritas em C/C++.
FN	Falso Negativo.
FP	Falso Positivo.
ImageJ	Software de domínio público para processamento de imagens desenvolvido pela Wayne Rasband National Institutes of Health.
MAPA	Ministério da Agricultura, Pecuária e Abastecimento.
MATLAB	MATrix LABoratory (software de alto desempenho para cálculo numérico e operações com matrizes).
OpenCV	<i>Open Computer Vision</i> (Biblioteca com rotinas para processamento de imagens e visão computacional, desenvolvida pela Intel).
ProEikon	Biblioteca para processamento de imagens e visão computacional de autoria do prof. Dr. Hae Yong Kim do Depto. de Eng. de Sistemas Eletrônicos da Escola Politécnica da Universidade de São Paulo.
RGB	<i>Red, Green e Blue</i> .
RNA	Rede Neural Artificial.
SE	Sistema Especialista
SIAD	Sistemas Inteligentes de Apoio à Decisão
SRPV	Sistema de Reconhecimento de Padrões Visuais.
SVC	Sistema de Visão Computacional.
SI	Sistema de Informação.
TI	Tecnologia da Informação.
VN	Verdadeiro Negativo.
VP	Verdadeiro Positivo.

LISTA DE SÍMBOLOS

$\nabla f(x, y)$	Vetor gradiente de uma imagem no ponto (x, y) .
$\ \nabla f(x, y)\ $	Magnitude do gradiente no ponto (x, y) .
$\theta(x, y)$	Direção do vetor gradiente no ponto (x, y) .
a	Imagem de material granular.
t_n	Núcleos (<i>kernels</i>) gerados para aplicação da granulometria.
c_i	Imagem resultante da correlação cruzada de a com t_n .
\tilde{t}	Correção pela média de t .
\bar{t}	Média de t .
$X = \{x_i, i=1, \dots, n\}$	Conjuntos de dados a serem agrupados pela técnica K-MEANS.
$C = \{c_k, k=1, \dots, k\}$	Conjuntos de <i>clusters</i> utilizados pela técnica K-MEANS.
u_k	Média dos <i>clusters</i> c_k .
$J(c_k)$	Resultado da aplicação da técnica K-MEANS.
l_{borda}	Limiar para a intensidade dos níveis de cinza das bordas.
$l_{composição}$	Limiar para a intensidade dos níveis de cinza da imagem com iluminação uniforme.
l_{corr}	Limiar para a taxa de correlação.
l_{sobrep}	Limiar para o valor de sobreposição.
r_{ij}	Raios das elipses.
s_i	Escala das elipses.
I_i	Imagens coloridas capturadas de amostras de feijão.
G_i	Imagens das amostras de feijão em níveis de cinza.
X_i	Magnitude do gradiente das imagens das amostras de feijão.
Z	Imagem da máxima magnitude do gradiente.
S	Imagem resultante da composição da imagem Z com a imagem G_0 .

SUMÁRIO

1. INTRODUÇÃO.....	15
1.1. Revisão da literatura	16
1.2. Objetivos	19
1.3. Justificativa e motivação	19
1.4. Estrutura da dissertação	20
2. FUNDAMENTAÇÃO TEÓRICA	21
2.1. Classificações do feijão brasileiro.....	21
2.2. Sistemas de informação e sistemas especialistas	22
2.3. Reconhecimento de padrões visuais	23
2.3.1. Sistemas de reconhecimento de padrões visuais	24
2.3.2. Processamento de imagens digitais	25
2.4. Técnicas de reconhecimento de padrões visuais empregadas no SE proposto ..	26
2.4.1. Detecção de bordas.....	27
2.4.2. Granulometria.....	28
2.5. Técnicas de inteligência computacional utilizadas na classificação de padrões.	30
2.5.1. Redes neurais artificiais	30
2.5.2. K-MEANS.....	33
3. MATERIAIS E MÉTODOS	35
3.1. Método de pesquisa.....	35
3.2. Delimitação dos experimentos e materiais empregados	35
3.3. Metodologia para a condução dos experimentos e parametrização do SE proposto	37
3.4. Avaliação de desempenho	40
4. RESULTADOS E DISCUSSÕES.....	42
4.1. Detalhamento do sistema especialista proposto.	42
4.2. Experimentos realizados com o SE proposto	55

5. CONCLUSÕES	63
5.1 SUGESTÕES PARA CONTINUIDADE DO TRABALHO	63
REFERÊNCIAS BIBLIOGRÁFICAS	64
ANEXO A: TRABALHO PUBLICADO EM ANAIS DE CONGRESSO.....	69
ANEXO B: TRABALHO PUBLICADO EM PERIÓDICO.	79

1. INTRODUÇÃO

A qualidade dos alimentos desempenha importante papel na indústria. Em muitos casos, propriedades visuais como cor, forma e tamanho são as principais características avaliadas pelos consumidores e são importantes fatores para a determinação do seu preço de mercado (FERNANDEZ; CASTILLERO; AGUILERA, 2005).

No caso do feijão comercializado no Brasil, sua qualidade está relacionada com a quantidade de defeitos e de misturas presentes nas amostras avaliadas e sua classificação é realizada com base em normas e procedimentos estabelecidos pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA, 2008).

Neste cenário, o processo de inspeção da qualidade visual dos alimentos precisa ser confiável, de modo que atenda às expectativas dos consumidores e às normas estabelecidas pelo governo. No entanto, é comum que esta inspeção ocorra de maneira manual, a qual pode demandar muito tempo, ser tediosa, gerar altos custos operacionais, ser passível a falhas humanas e apresentar dificuldades para padronizar os resultados, tornando desejável a sua automatização (BROSNAN; SUN, 2004; KILIC et al., 2007; PATIL; MALEMATH; YADAHALLI, 2011).

Os Sistemas de Reconhecimento de Padrões Visuais (SRPV), também conhecidos como Sistemas de Visão Computacional (SVC), são empregados para tornar possível esta automatização. Tais sistemas são utilizados no controle de qualidade com o objetivo de identificar defeitos nos alimentos durante o processo industrial e de classificá-los de acordo com suas características externas (GONZALEZ; LINUESA; GARCIA, 2001; LEEMANS; DESTAIN, 2004). Outro ganho que estes sistemas proporcionam é a possibilidade de avaliar a qualidade visual dos alimentos de maneira não destrutiva (TIMMERMANS, 1998).

A automatização de processos pode ser associada aos Sistemas Especialistas (SE), ou Sistemas Inteligentes de Apoio à Decisão (SIAD), tendo em vista que estes consistem em sistemas computacionais (envolvendo hardware e software) que tentam simular os seres humanos na execução de tarefas. (LAUDON; LAUDON, 2010).

Neste contexto, este trabalho consiste no desenvolvimento de um SE aplicado à inspeção da qualidade visual de grãos de feijão, capaz de classificá-los de acordo com a coloração de suas películas de maneira robusta, visando proporcionar maior agilidade e confiabilidade neste processo. Para tanto, foram empregadas técnicas de reconhecimento de padrões visuais e de inteligência computacional. Além disso, foi proposto um método para detecção dos grãos que permitiu superar uma limitação presente nos SRPV propostos na

literatura para análise e classificação de sementes e grãos, a qual dificulta a aplicabilidade de tais sistemas em processos industriais.

O método proposto para detecção de grãos inclui as etapas de aquisição, pré-processamento e segmentação, descritas nas seções 4.1.1 a 4.1.3.

Destaca-se que a classificação proposta neste trabalho não abrange defeitos presentes em feijões, e sim suas classes, determinadas pela coloração de suas películas. A classificação do feijão a partir de suas classes é uma importante etapa no processo de inspeção da sua qualidade, visto que falhas causadas neste processo acarretam no não atendimento às normas do MAPA, podendo ter como consequência, penalizações com o pagamento de multas. Assim, um processo mais acurado pode resultar, inclusive, em um diferencial competitivo para empresa, uma vez que oferece menor probabilidade de erros no processo de inspeção.

1.1. Revisão da literatura

Diversos trabalhos voltados para a inspeção automática da qualidade visual de alimentos estão disponíveis na literatura. Em geral, estes trabalhos empregam técnicas de reconhecimento de padrões visuais para atingirem seus objetivos.

Paulus e Schrevens (1999) desenvolveram um algoritmo que utiliza a expansão de Fourier para caracterizar os formatos de maçãs, possibilitando distinguir diferentes tipos. A abordagem matemática realizada demonstrou serem necessárias quatro imagens de cada maçã para quantificar a sua forma média. Os resultados mostraram a viabilidade do algoritmo proposto ao distinguir diferentes cultivares de maçãs a partir de amostras aleatórias.

Riyadi et al. (2007) propuseram um SVC destinado para a classificação do tamanho do mamão papaya. Nesse trabalho foi utilizado o método proposto por Otsu (1979) para binarização e operadores morfológicos para tratamento da imagem, de modo que o fruto fosse destacado do fundo da imagem. Além disso, características como área, diâmetro médio e perímetro foram utilizadas para treinamento e testes de redes neurais artificiais, resultando em uma precisão superior a 94% no processo de classificação.

Savakar e Anami (2009) apresentaram em seu trabalho um SVC que emprega redes neurais artificiais para a identificação e classificação de grãos, frutas e flores a partir da extração de cores, texturas e características morfológicas. Os resultados mostraram que a combinação destas características proporcionou uma média de acurácia de 94,1% para grãos, como milho, arroz, trigo entre outros, 84% para cinco tipos de mangas e 90,1% para dez tipos de flores jasmim.

Marco, Reis e Saraiva (2009) propuseram uma metodologia para avaliação *off-line* (que não ocorre em tempo real) da qualidade da cobertura aplicada em flocos de cereais, composta por corantes e aromatizantes. Esta avaliação ocorreu a partir do emprego da técnica de classificação supervisionada, para identificar regiões não uniformes da cobertura aplicada nos flocos analisados. Os resultados obtidos se mostraram consistentes com a avaliação realizada por especialistas e dessa forma, indicou que o método proposto é capaz de classificar diferentes amostras de flocos de cereais a partir de classes pré-determinadas de qualidade.

Rodríguez-Pulido et al. (2013) empregaram técnicas de visão computacional para estimar a composição de cores em alimentos como frutas e vegetais, com diferentes tamanhos, cores e texturas. Foi utilizado um método de agrupamento (*clustering*) subtrativo para agrupar os pixels das imagens com base na sua coloração e os grupos formados foram destacados por elipses. Os resultados deste trabalho foram positivos, de modo que possibilita a aplicação de técnicas de estatística multivariada para a classificação e previsão de propriedades químicas em alimentos.

Cheng e Ying (2004) desenvolveram um sistema para inspeção da qualidade visual de grãos de arroz baseada na transformada de Hough. Cinco tipos de grãos foram considerados para a tarefa de inspeção, que consistiu em inspecionar as glumas dos grãos. Os resultados apontaram precisão média de 96% para os grãos normais, 92% para grãos com pequenas fissuras e 87% para grãos com glumas não fechadas por completo.

Aggarwal e Mohan (2010) também avaliaram a qualidade visual de grãos de arroz a partir da distribuição da relação de aspecto. Neste trabalho o software *ImageJ* foi utilizado para realizar o pré-processamento e a segmentação das imagens e o software MATLAB foi utilizado para a extração e a representação dos dados. Três tipos de grãos, denominados de completo (grupo 1), meio (grupo 2) e quebrado (grupo 3) foram considerados para a classificação, na qual o objetivo foi verificar a extensão de *offsize* (tamanho inadequado) presentes em cada grupo comercializado, sendo que os resultados indicaram um maior *offsize* no grupo 2. Também foi verificada a capacidade do método proposto em quantificar a composição de mistura dos tipos de arroz presentes nas amostras, em que foi obtida precisão com margem de erro de 10%.

Especificamente para a tarefa de classificação automática de grãos de feijão, apenas os trabalhos de Kiliç et al. (2007), Venora et al. (2009), Laurent et al. (2010) e Pires (2012) foram encontrados na revisão da literatura, realizada a partir das palavras chave “*beans*”, “*Image processing*”, “*classification system*”, “*classification of food*”, “*automatic*

classification”, “*automatic inspection*”, “*inspection of food*”, em inglês e português, nas bases CAPES, Scielo, ScienceDirect, IEEE Xplore e Google Scholar.

Kiliç et al. (2007) propuseram um SVC destinado para a classificação de grãos de feijão baseado na sua coloração e no seu tamanho. O software MATLAB foi utilizado na codificação das técnicas empregadas, como segmentação, binarização, detecção das bordas e quantificação de cores. Para a tarefa de classificação, na qual se obteve como resultado uma taxa de acertos de 90,6%, empregou-se uma rede neural artificial supervisionada.

Venora et al. (2009) utilizaram técnicas de análise e processamento de imagens e classificadores estatísticos para identificar 15 variedades de feijão cultivados em duas regiões da Itália. Características como forma, cor e tamanho dos grãos foram consideradas e os resultados obtidos foram de 98,49% para a taxa de acertos. Além disso, o método proposto se mostrou capaz de diferenciar o feijão pelo seu ano de cultivo ou pela sua região de cultivo.

Laurent et al. (2010) avaliaram a relação entre alterações na coloração do feijão durante o armazenamento com o fenômeno conhecido como “grãos difíceis de cozinhar”. Para tanto, utilizaram histogramas de cor e análises estatísticas. Os resultados constataram a existência desta relação e confirmaram a eficiência do modelo proposto a partir do processamento de imagens digitais.

Pires (2012) propôs um protótipo para auxiliar na tarefa de inspeção da qualidade visual de grãos de feijão. Um software foi desenvolvido com o objetivo detectar grãos pretos e mulatos contidos em amostras de feijão carioca. Neste software foi empregado o método de OTSU (OTSU, 1979) para limiarização das imagens e algoritmos baseados em morfologia matemática para identificação dos centros de cada grão de feijão. Os resultados apresentaram taxas de acertos de 91,7% e taxa de precisão de 94,2%.

Os quatro trabalhos destinados especificamente para a classificação de grãos de feijão demonstram a necessidade e a importância de processos automáticos para a inspeção da sua qualidade visual, porém os trabalhos de Kiliç et al. (2007), Venora et al. (2009) e Laurent et al. (2010) apresentam uma limitação com relação a disposição das amostras, em que os grãos precisam estar separados uns dos outros para a sua correta detecção e classificação. Além disso, estes trabalhos não contemplam os grãos de feijão mais comercializados no Brasil. Já no trabalho proposto por Pires (2012), são considerados os principais grãos de feijão comercializados no Brasil. Contudo, conforme apontado pelo próprio autor, os métodos empregados para a segmentação dos grãos são muito sensíveis à presença de sombras, variações na iluminação e na tonalidade das películas, o que pode gerar altos números de falso positivos e falso negativos. Não obstante, o método proposto não é capaz de discriminar as

classes de grãos presentes na amostra, o que também pode limitar a sua aplicação prática na indústria de alimentos.

1.2. Objetivos

Este trabalho teve como principal objetivo desenvolver um SE aplicado à inspeção da qualidade visual de grãos de feijão, capaz de classificá-los de acordo com a coloração de suas películas, de modo que apresentasse altas taxas de acertos, precisão e acurácia, além de superar a limitação relacionada com a disposição dos grãos na amostra a ser analisada, presente nos SRPV propostos na literatura para a mesma finalidade. Para tanto, os seguintes objetivos específicos foram considerados:

- Investigar as limitações existentes nos SRPV voltados para inspeção da qualidade visual de grãos;
- Desenvolver um SE baseado em técnicas de reconhecimento de padrões visuais e de inteligência computacional, capaz de analisar e classificar grãos de feijão a partir de imagens digitais;
- Realizar experimentos para aferir o desempenho do SE proposto, a partir das taxas de acertos, precisão e acurácia, além das analisar suas limitações e sua aplicabilidade na indústria de alimentos;

1.3. Justificativa e motivação

Devido ao aumento da expectativa dos consumidores por maior qualidade nos alimentos comercializados e normas de segurança mais rígidas, cresce a necessidade por sistemas de inspeção da qualidade visual mais precisos e rápidos (BROSNAN; SUN, 2004; PATEL et al., 2012).

Com os recentes avanços em hardware e software, novos estudos destinados à automatização do processo de inspeção da qualidade visual têm sido realizados para a indústria de alimentos, com o objetivo de gerar sistemas de apoio à decisão de menores custos e que gerem bons resultados (LOCHT; THOMSEN; MIKKELSEN, 1997; SUN, 2000). Porém, ainda há desafios a serem superados pelos SRPV voltados para inspeção automática da qualidade visual de alimentos, inclusive para a qualidade dos feijões, o que impulsiona este trabalho.

Um dos principais desafios está em possibilitar que os grãos a serem analisados sejam dispostos de maneira similar ao cenário encontrado na indústria de alimentos, ou seja, sem a necessidade de mantê-los separados para facilitar a sua identificação, conforme pode ser visto nos trabalhos de Kiliç et al. (2007), Venora et al. (2009), Laurent et al. (2010) e Aggarwal e Mohan (2010).

Não obstante, uma alta taxa de acertos também é imprescindível para tais sistemas e melhorias neste aspecto também são necessárias. No trabalho de Pires (2012), por exemplo, a questão relacionada a não necessidade de distanciamento dos grãos é parcialmente resolvida, porém os resultados obtidos teriam de ser melhorados para tornar viável a aplicação do sistema proposto em situações práticas.

Por essas razões, torna-se relevante a realização deste trabalho, uma vez os objetivos propostos visam superar algumas limitações existentes nos SRPV voltados para a inspeção da qualidade visual de grãos de feijão, de forma que seja possível a sua aplicação prática em processos industriais.

1.4. Estrutura da dissertação

Esta dissertação está estruturada do seguinte modo: no segundo capítulo há uma fundamentação teórica a respeito dos conceitos relacionados ao feijão e suas classificações, conceitos acerca dos sistemas de informação e sistemas especialistas, além da descrição das principais técnicas de reconhecimento de padrões visuais em imagens digitais e de inteligência computacional empregadas neste trabalho. No terceiro capítulo são descritos os materiais utilizados, além do método de condução dos experimentos e de avaliação dos resultados. No quarto capítulo são apresentados e discutidos os resultados obtidos. Por fim, no quinto capítulo, apresentam-se as conclusões e as propostas de continuidade deste trabalho.

2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são descritos os conceitos relacionados ao feijão e suas classificações, sistemas de informação e sistemas especialistas e as técnicas de reconhecimento de padrões visuais e de inteligência computacional empregadas no desenvolvimento do SE proposto.

2.1. Classificações do feijão brasileiro

O feijão é classificado por grupos, classes e tipos, de acordo com o Regulamento Técnico do Feijão, estabelecido pelo Ministério da Agricultura, Pecuária e Abastecimento, através da Instrução Normativa nº 12 de 28/03/2008 (MAPA, 2008). Com base neste regulamento, os grãos provenientes da espécie *Phaseolus vulgaris L.* pertencem ao Grupo I (Feijão Comum) e os grãos provenientes da espécie *Vigna unguiculata (L) Walp* pertencem ao grupo II (Feijão-Caupi).

Para cada grupo há quatro classes de feijão que são determinadas com base na coloração da sua película. Essas classes são definidas como: branco, preto, cores e misturado e a sua classificação e comercialização devem respeitar normas que estabelecem os limites de misturas de cores toleráveis em cada classe, conforme apresentado na Tabela 1, adaptado de (MAPA, 2008).

Tabela 1: Percentual de mistura tolerada em cada classe de feijão

Grupo	Classe	Percentual mínimo de grãos da classe	Observações
I	Branco	97%	-
I	Preto	97%	-
I	Cores	97%	-
I	Misturado	-	Produto que não atende às especificações de nenhuma das classes anteriores.
II	Branco	90%	-
II	Preto	90%	-
II	Cores	90%	-
II	Misturado	-	Produto que não atende às especificações de nenhuma das classes anteriores.

Fonte: Adaptado de MAPA (2008).

Os tipos estão relacionados com os percentuais de tolerância de defeitos previstos em cada grupo. Neste caso, cada grupo de feijão é classificado em três tipos ou enquadrado como “fora de tipo” ou “desclassificado”. O tipo 1 apresenta um percentual de defeitos menor do que os tipos 2 e 3. O “fora de tipo” representa o feijão que não atende aos limites de tolerância

do tipo 3, mas que ainda assim pode ser comercializado. O “desclassificado” representa o feijão que possui um percentual de defeitos graves ou insetos mortos acima dos limites de tolerância estabelecidos para o “fora de tipo”, e dessa forma, não pode ser comercializado. (MAPA, 2008).

Dentre os defeitos presentes no feijão estão os grãos avariados, ardidados, mofados, brotados, enrugados, manchados, descoloridos, amassados, partidos (bandas), quebrados (pedaços), carunchados, danificados por insetos e matérias estranhas (MAPA, 2008).

O SE desenvolvido neste trabalho classifica os feijões considerando apenas a classe. Não há a classificação por defeitos. Dessa forma, a sua aplicação dá-se após a remoção dos grãos defeituosos e das matérias estranhas, que é uma das etapas iniciais do processo de inspeção da qualidade visual de grãos de feijão.

Na seção seguinte são abordados os conceitos teóricos acerca dos sistemas de informação e sistemas especialistas, que forneceram subsídios para o desenvolvimento deste trabalho.

2.2. Sistemas de informação e sistemas especialistas

Os Sistemas de Informação (SI) podem ser definidos como um conjunto integrado de pessoas, recursos tecnológicos e recursos de comunicação, voltados para a coleta, transformação e disseminação da informação em uma organização (O’BRIEN, 2010).

Segundo O’Brien (2010), os SI podem desempenhar três importantes contribuições para uma empresa:

- Suporte de seus processos e operações;
- Suporte na tomada de decisões de seus colaboradores nos mais diversos níveis hierárquicos;
- Suporte em suas estratégias em busca de vantagem competitiva.

De acordo com Beal (2004), o emprego dos SI, apoiados pela Tecnologia da Informação (TI), tem proporcionado informações de melhor qualidade, resultando em um significativo aumento da produtividade nas áreas gerenciais e operacionais das organizações.

Neste cenário, os Sistemas Especialistas (SE), ou Sistemas Inteligentes de Apoio à Decisão (SIAD), são sistemas de informação alimentados com a expertise de um especialista,

com o objetivo de realizar tarefas em um domínio específico do conhecimento, ou de apoiar na tomada de decisão (LAUDON; LAUDON, 2010; O'BRIEN, 2010).

Os SE podem ser aplicados em atividades como: gerenciamento de decisões, diagnóstico de problemas de operação, projeto/configuração, seleção/classificação e monitoração/control de processos (O'BRIEN, 2010).

A sua arquitetura é composta por uma Base de Conhecimento (BC), que armazena conhecimentos e fatos acerca de uma área específica, além de um conjunto de métodos de inteligência computacional para manipulação da BC, denominados mecanismos de inferência (O'BRIEN, 2010; ROSARIO et al., 2011). Trabalhos que utilizam SE em processos industriais podem ser vistos em Baptista (2004), Souza et al. (2004), Rosario et al. (2011) e Malinowski (2012).

O sistema desenvolvido neste trabalho pode ser caracterizado como um SE, uma vez que ele possui uma BC que contém dados e parâmetros específicos sobre o feijão, além de um conjunto de regras de inferência, obtidas a partir do uso de técnicas de inteligência computacional e de reconhecimento de padrões visuais, para possibilitar a adequada detecção e classificação dos grãos.

Nas próximas seções, são abordados conceitos relacionados com reconhecimento de padrões visuais e inteligência computacional, uma vez que técnicas oriundas destas áreas são empregadas no desenvolvimento do SE proposto neste trabalho.

2.3. Reconhecimento de padrões visuais

Um padrão pode ser definido como uma entidade que possua algum tipo de característica numérica ou simbólica que possa ser extraída, ou ainda como uma amostra que possa representar uma classe ou tipo (OSÓRIO, 1991; ARAÚJO, 2009).

O reconhecimento de padrões está relacionado com a identificação de um elemento a partir da comparação de suas características extraídas com algum tipo de conhecimento pré-existente. Seu estudo pode ser dividido em estudos direcionados para a compreensão das habilidades desenvolvidas pelos seres humanos e outros organismos vivos e estudos voltados para a construção de máquinas que apresentem a capacidade de reconhecer padrões de forma semelhante aos seres humanos (TOU; GONZÁLES, 1981).

Problemas relacionados ao reconhecimento automático, descrição, classificação e agrupamento de padrões estão presentes em diversas áreas como biologia, psicologia,

medicina, marketing, visão computacional, inteligência artificial e sensoriamento remoto (JAIN; DUIN; MAO, 2000).

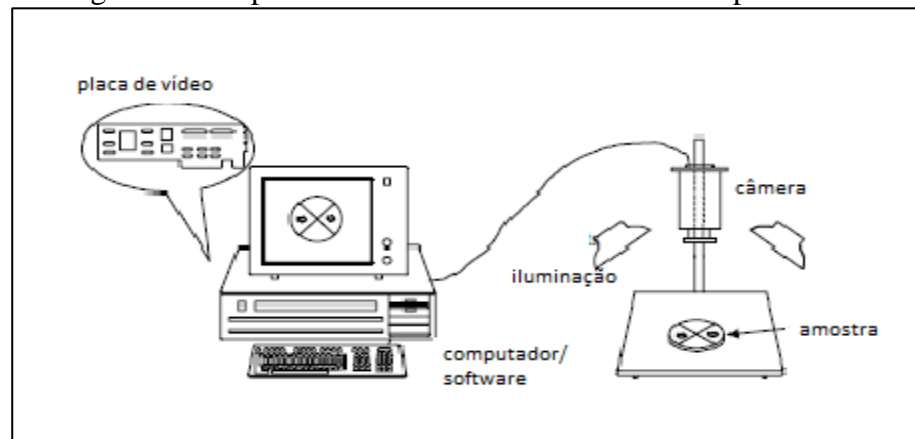
Na área de visão computacional, explora-se o desenvolvimento de sistemas de reconhecimento de padrões visuais (SRPV) ou sistemas de visão computacional (SVC), com o objetivo de simular as capacidades da visão humana. Tais sistemas são conceituados a seguir.

2.3.1. Sistemas de reconhecimento de padrões visuais

Os SRPV ou SVC consistem de métodos e técnicas que possibilitem a interpretação de imagens a partir de sistemas computacionais (CONCI; AZEVEDO; LETA, 2008). Neste contexto, o reconhecimento de padrões na área de visão computacional tem por objetivo detectar padrões em imagens digitais com base em modelos, máscaras ou *templates* de referência (ARAÚJO, 2009).

No geral, os SRPV são constituídos por cinco componentes (Figura 1): recursos de iluminação, dispositivo para captura de imagens, placa de vídeo, equipamentos de hardware e software (WANG; SUN, 2002).

Figura 1: Componentes de um sistema de visão computacional.



Fonte: Adaptado de Wang e Sun (2002).

O sistema de iluminação possibilita reduzir reflexos, sombras e ruídos na imagem a ser analisada, influenciando diretamente no tempo de processamento, eficiência e precisão do sistema (NOVINI, 1995; GUNASEKARAN, 1996). Diferentes tipos de lâmpadas, como fluorescente, dicroicas, halógenas e lâmpadas LED são empregados nos SRPV (ALFATNI et al., 2011).

Para a captura de imagens, dispositivos como ultrassom, raio-x e câmeras monocromáticas e coloridas são utilizados (BROSNAN; SUN, 2004).

A placa de vídeo atua no processo de digitalização da imagem e a sua escolha depende de fatores como: o tipo da saída do dispositivo de captura, resolução desejada para a imagem a ser processada, além da capacidade de processamento do computador utilizado (BROSNAN; SUN, 2004).

Os equipamentos de hardware e software possibilitam a integração entre o sistema de iluminação, o dispositivo para a captura das imagens e a placa de vídeo. Além disso, também possibilitam o processamento da imagem digitalizada, a partir das linguagens de programação (WANG; SUN, 2002).

A seguir, são detalhados os conceitos relacionados ao processamento de imagens digitais, essenciais para a construção de um SRPV.

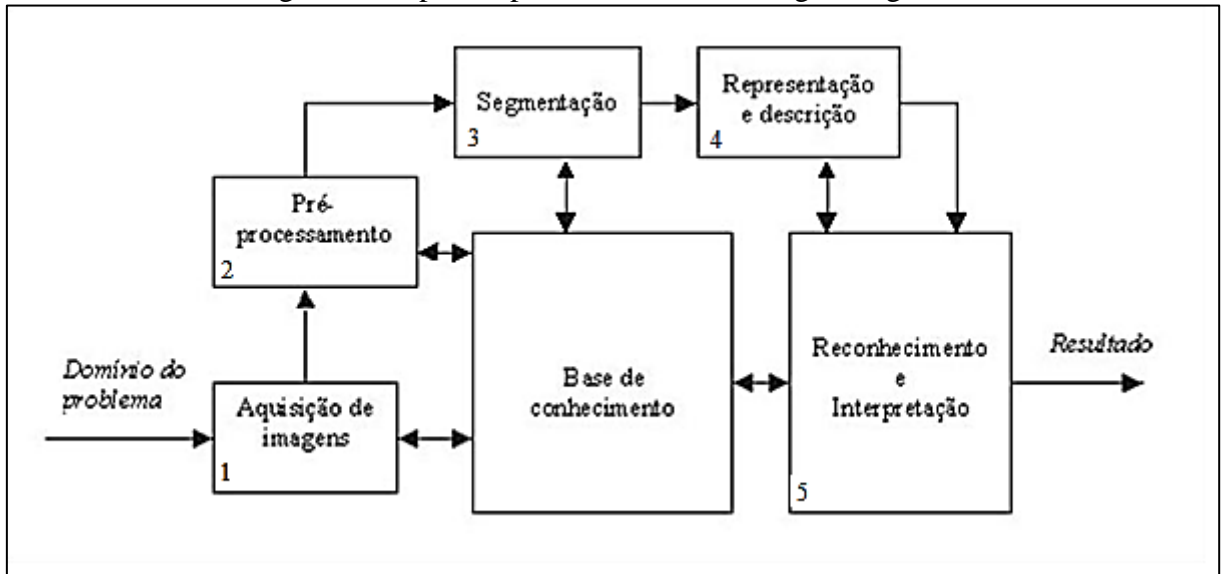
2.3.2. Processamento de imagens digitais

Uma imagem digital pode ser definida como uma função $f(x,y)$ discretizada a partir dos processos de amostragem e quantização. A amostragem discretiza o domínio de definição da imagem, onde x e $y \in \mathfrak{R}^2$ e representam as coordenadas espaciais em uma matriz e a quantização consiste na definição da intensidade do nível de cinza presente em cada ponto da função, denominado pixel (acrônimo do inglês *picture element*) (GONZALEZ; WOODS, 2002; CONCI; AZEVEDO; LETA, 2008; PEDRINI; SCHWARTZ, 2008).

O processamento de imagens digitais compreende o uso de técnicas para a captura, representação e transformação de imagens, com o objetivo de extrair e identificar informações e melhorar a qualidade visual da imagem, auxiliando na interpretação humana e na interpretação automática através de máquinas (PEDRINI; SCHWARTZ, 2008).

As seguintes etapas compõe o processamento de imagens (Figura 2): aquisição de imagens, pré-processamento, segmentação, representação e descrição e reconhecimento e interpretação (GONZALEZ; WOODS, 2002; PEDRINI; SCHWARTZ, 2008).

Figura 2: Etapas do processamento de imagens digitais.



Fonte: Adaptado de Gonzalez e Woods (2002).

Etapa 1: A aquisição de imagens consiste na captura da imagem por meio de um dispositivo ou sensor; Etapa 2: O pré-processamento tem por objetivo melhorar a qualidade da imagem através de técnicas para redução de ruídos, ajuste de contraste ou brilho e suavização, tornando o processamento da imagem mais eficiente; Etapa 3: A segmentação visa extrair ou isolar a região de interesse da imagem a ser processada; Etapa 4: A etapa de representação e descrição se refere ao processo de extração das características que representam os padrões; Etapa 5: A última etapa envolve o reconhecimento e a interpretação, na qual o reconhecimento está relacionado com a atribuição de um rótulo aos padrões localizados na imagem a partir dos atributos que os descrevem e a interpretação está relacionada com a atribuição de um significado ao conjunto de padrões reconhecidos. A base de conhecimento fornece dados relacionados ao domínio do problema com o intuito de auxiliar na execução de todas as etapas (GONZALEZ; WOODS, 2002; PEDRINI; SCHWARTZ, 2008).

2.4. Técnicas de reconhecimento de padrões visuais empregadas no SE proposto

Nas subseções seguintes são descritas algumas técnicas de reconhecimento de padrões em imagens digitais, empregadas neste trabalho.

2.4.1. Detecção de bordas

A borda pode representar o limite de um objeto ou a fronteira entre duas regiões com diferentes intensidades de níveis de cinza. O processo de detecção de bordas consiste na identificação de mudanças significativas dos níveis de cinza em regiões locais de uma imagem (PEDRINI; SCHWARTZ, 2008). Essas mudanças podem ser descritas através do conceito de derivadas, sendo o operador *gradiente* um dos mais utilizados.

O vetor gradiente $\nabla f(x, y)$ de uma imagem no ponto (x, y) é definido pela equação 1:

$$\nabla f(x, y) = \left[\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right], \quad (1)$$

Na qual $\partial f(x, y)/\partial x$ e $\partial f(x, y)/\partial y$ representam as derivadas parciais discretas nas direções x e y , e podem ser definidas pelas equações 2 e 3, respectivamente:

$$\frac{\partial f(x, y)}{\partial x} \approx f(x + 1, y) - f(x, y), \quad (2)$$

$$\frac{\partial f(x, y)}{\partial y} \approx f(x, y + 1) - f(x, y). \quad (3)$$

A magnitude do gradiente pode ser obtida através da equação 4 e representa a maior taxa de variação de $f(x, y)$, por unidade de distância na direção ∇f .

$$\|\nabla f(x, y)\| = \sqrt{\frac{\partial f(x, y)}{\partial x}^2 + \frac{\partial f(x, y)}{\partial y}^2}. \quad (4)$$

A direção do vetor gradiente pode ser obtida pela equação 5:

$$\theta(x, y) = \tan^{-1} \left(\frac{\partial f(x, y)}{\partial y} / \frac{\partial f(x, y)}{\partial x} \right). \quad (5)$$

Definindo um limiar lm , é possível determinar se um ponto pertence à borda quando $\|\nabla f(x, y)\| \geq lm$.

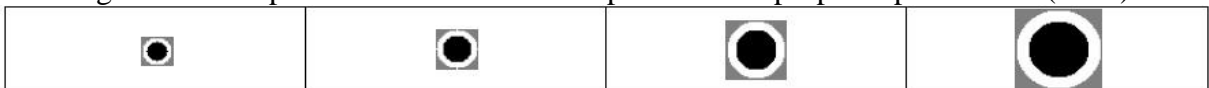
2.4.2. Granulometria

A granulometria é o processo de estimação da distribuição dos tamanhos de objetos/grãos contidos em imagens que possuam materiais granulares (PEDRINI; SCHWARTZ, 2008; MARUTA, 2011). Seu objetivo primário é obter o espectro de padrões (*pattern spectrum*) de uma imagem, onde espectro de padrões se refere ao histograma de objetos em função do tamanho, e seu objetivo secundário é localizar espacialmente cada grão e suas propriedades, como tamanho, forma, contraste etc. (MARUTA, 2011).

Há dois grupos principais de algoritmos aplicados a granulometria: algoritmos baseados em morfologia matemática e algoritmos baseados na detecção de contornos. No primeiro caso, de acordo com Maruta (2011), os resultados não são satisfatórios por não serem intuitivos e por não gerarem espectros que indiquem a quantidade dos grãos e suas propriedades. No segundo caso, os resultados também podem ser insatisfatórios dependendo da aplicação desejada, devido a sua sensibilidade a ruídos durante o pré-processamento, conforme exposto por Maruta (2011).

Maruta (2011) propôs um método para aplicação da granulometria baseada na correlação cruzada. Neste método, a obtenção do espectro de padrões de uma imagem de material granular a é obtida através da correlação cruzada de a com núcleos (*kernels*) circulares (ou outros formatos como, por exemplo, elipses) t_1, t_2, \dots, t_n com diferentes raios (Figura 3).

Figura 3: Exemplos de núcleos utilizados pelo método proposto por Maruta (2011).



Fonte: Maruta (2011).

O resultado da correlação cruzada de a com os núcleos circulares t_1, t_2, \dots, t_n , denotado por c_1, c_2, \dots, c_n é obtido pela equação 6:

$$c_i(x, y) = \tilde{t}_i(x, y) \circ a(x, y). \quad (6)$$

Na qual \tilde{t} representa a correção pela média de t , definida pela equação 7:

$$\tilde{t}(x, y) = t(x, y) - \bar{t} \text{ e } \bar{t} = \frac{1}{MN} \sum_{x=1}^m \sum_{y=1}^n t(x, y). \quad (7)$$

A máxima correlação, denotada por C , é obtida através da equação 8. Nesta equação, um pico em $C(x, y)$ corresponde a um objeto aproximadamente circular em a .

$$C(x, y) = \text{MAX}_{i=1}^n [c_i(x, y)]. \quad (8)$$

O argumento i da correlação máxima M (equação 9) possibilita obter a escala do objeto detectado.

$$M(x, y) = \text{ARGMAX}_{i=1}^n [c_i(x, y)]. \quad (9)$$

A partir das equações 6 a 9, é possível descartar da imagem resultante C os objetos com baixa correlação ou que apresentem alguma sobreposição significativa com alguns objetos de maior correlação. Para descartar os objetos com baixa correlação, é utilizada a variável l_{corr} , que pode receber valores entre 0 e 1 para definir a exclusão. Neste caso, quando $C(x, y) < l_{corr}$, o objeto é descartado. Para definir o valor máximo de sobreposição permitida, a variável l_{sobrep} é utilizada, recebendo valores entre 0 e 1. Objetos que apresentem alguma sobreposição maior do que l_{sobrep} com algum outro objeto que possua maior correlação também são descartados. Com base nos experimentos apresentados em Maruta (2011), este método é robusto a ruídos e pode detectar objetos com pouquíssimo contraste.

Neste trabalho, emprega-se este método de granulometria para detecção dos grãos de feijão, trocando os *kernels* circulares por elipses e empregando uma etapa de pré-processamento que potencializa os resultados da detecção.

2.5. Técnicas de inteligência computacional utilizadas na classificação de padrões

Os classificadores, também denominados algoritmos ou técnicas de classificação, visam estabelecer uma relação entre as propriedades extraídas de um objeto ou amostra, com um conjunto de rótulos, resultando em classes compostas por objetos que compartilham de propriedades em comum (PEDRINI; SCHWARTZ, 2008).

Neste trabalho, duas das principais técnicas de inteligência computacional aplicadas em tarefas de classificação de padrões foram utilizadas: Redes Neurais Artificiais e K-MEANS, as quais são detalhadas a seguir.

2.5.1. Redes neurais artificiais

As redes neurais artificiais (RNA) consistem em uma técnica de inteligência computacional que têm como inspiração a estrutura e o funcionamento do cérebro humano. Assim como no cérebro, as RNA são compostas de elementos individuais simples, chamados de neurônios, que se interconectam, formando redes capazes de armazenar e transmitir a informação provinda do exterior ou de outros neurônios (KASABOV, 1998).

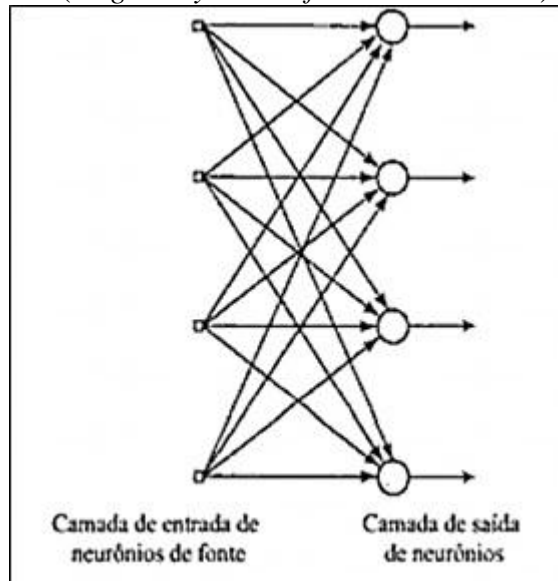
Dentre suas características, destaca-se a capacidade de aprendizado a partir de exemplos e a possibilidade de generalizar informações (PEDRINI; SCHWARTZ, 2008).

Há diversos modelos de RNA, dos quais cada um é destinado para a resolução de uma determinada classe de problema, como reconhecimento de padrões e classificação, processamento de imagens e visão computacional e processamento de sinais (TATIBANA; KAETSU, 2006).

A arquitetura de rede é uma das características que diferenciam cada modelo de RNA e está relacionada com a maneira pela qual os neurônios de uma rede neural estão estruturados, podendo ser classificada em três classes específicas (HAYKIN, 2001): redes alimentadas diretamente com camada única (*Single-Layer Feedforward Networks*), redes alimentadas diretamente com múltiplas camadas (*Multi-Layer Feedforward Networks*) e redes recorrentes (*Feedback Networks*).

As redes alimentadas diretamente com camada única (Figura 4) são caracterizadas por possuir somente as camadas de entrada e saída. O sinal percorre a rede em um único sentido, sem realimentação e sem a presença de camadas ocultas.

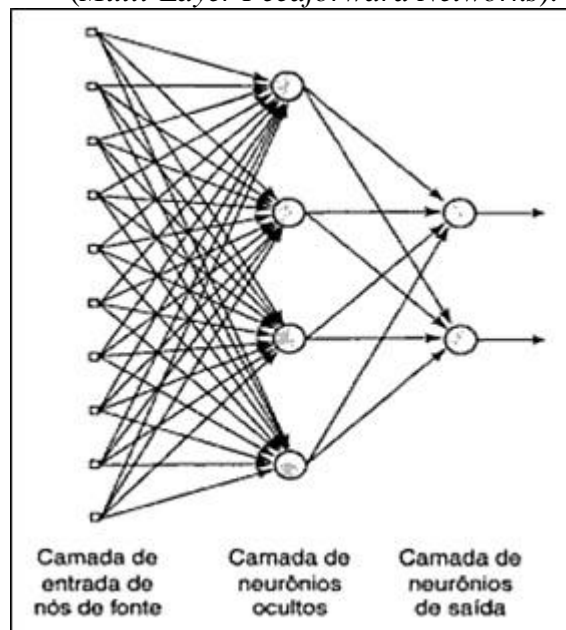
Figura 4: Redes alimentadas diretamente com camada única
(*Single-Layer Feedforward Networks*).



Fonte: Haykin (2001).

As redes alimentadas diretamente com múltiplas camadas (Figura 5) também não possuem realimentação e a conexão entre os neurônios é unidirecional, porém difere da anterior por possuir camadas ocultas.

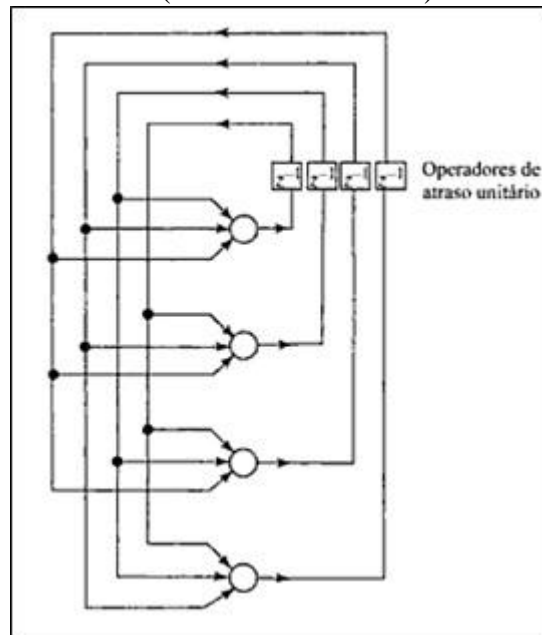
Figura 5: Redes alimentadas diretamente com múltiplas camadas
(*Multi-Layer Feedforward Networks*).



Fonte: Haykin (2001).

As redes recorrentes (Figura 6) são caracterizadas por possuir a realimentação, em que a saída de um neurônio pode alimentar neurônios da mesma camada ou de camadas anteriores (JAIN; MAO, 1996; HAYKIN, 2001).

Figura 6: Redes recorrentes
(Feedback Networks).



Fonte: Haykin (2001).

O aprendizado de uma RNA é obtido a partir da etapa de treinamento, que consiste do processo de atualização iterativa dos pesos até que seja encontrada uma função capaz de classificar corretamente todas as amostras contidas no conjunto de treinamento (PEDRINI; SCHWARTZ, 2008). Este aprendizado pode ser classificado em aprendizado supervisionado e aprendizado não supervisionado.

O aprendizado supervisionado é caracterizado pelo processo de supervisão dos resultados de saída da rede, que são comparados com os resultados esperados para os padrões das classes submetidas para treinamento. *Multi-Layer Perceptrons* é um exemplo de modelo de aprendizado supervisionado (OSÓRIO, 1991).

O aprendizado não supervisionado (auto-organizável) não possui um agente externo para indicar os resultados desejados para os padrões de entrada. Neste caso, a rede identifica padrões com base na redundância dos dados de entrada e com base em padrões semelhantes aos esperados pela rede. Os Mapas Auto-Organizáveis, ou redes de Kohonen são um exemplo de um modelo de RNA de aprendizado não supervisionado (ARAÚJO; BONALDO; SANTOS, 2004).

As RNA têm sido utilizadas com sucesso nos SRPV voltados para a indústria de alimentos (YING et al., 2003). Aplicações de RNA voltadas para a classificação e inspeção da qualidade visual de alimentos podem ser vistas em KILIÇ et al. (2007); LIU et al. (2011); GUZMAN e PERALTA (2008) e SAVAKAR (2012).

2.5.2. K-MEANS

Problemas relacionados à *clustering* estão presentes em diversos domínios de aplicação, como Redes Neurais Artificiais, Inteligência Artificial e Estatística (ALSABTI; RANKA; SINGH, 1997). Seu processo consiste em particionar ou agrupar conjuntos de padrões em *clusters* disjuntos e a sua aplicação pode ser dividida em dois grupos: algoritmos hierárquicos e algoritmos particionais (LIU; YU, 2009).

Os algoritmos hierárquicos podem ser aglomerativos, nos quais cada elemento inicia representando um *cluster* e que ao longo do processo se une a outros *clusters* similares, ou podem ser divisivos, onde todos os elementos iniciam pertencendo a um único grupo, que se subdivide recursivamente em *clusters* menores (PEDRINI; SCHWARTZ, 2008; LIU; YU, 2009).

Os algoritmos particionais caracterizam-se por não possuírem uma estrutura hierárquica, gerando apenas um único particionamento. Neste caso, é necessário determinar previamente a quantidade de *clusters* (k) que serão gerados pelo particionamento. Dentre os algoritmos particionais, os algoritmos baseados na otimização de funções de custo visam particionar os elementos presentes em um espaço de características em k *clusters*, previamente definidos, em que cada *cluster* é representado por um centróide. A função de otimização consiste na minimização ou maximização de algum critério específico com base nos centróides gerados (KANUNGO et al., 2002; PEDRINI; SCHWARTZ, 2008).

Neste contexto, o K-MEANS é o algoritmo particional mais popular utilizado em problemas de *clustering*. Sua popularidade se deve principalmente pela facilidade para implementação, simplicidade e eficiência (LIU; YU, 2009). As principais etapas do algoritmo K-MEANS são:

1. Definir a quantidade de *clusters*.
2. Escolher os elementos do conjunto que representarão os *clusters* iniciais. Esses elementos são chamados de centróides.

3. Associar cada elemento do conjunto ao centróide mais próximo, com base em algum critério de similaridade ou distância.
4. Após todos os elementos estarem associados a algum centróide, recalcular os centróides de cada *cluster*.
5. Repetir os passos 3 e 4 até que não ocorra mais variações nos centróides de cada *cluster* gerado.

O algoritmo K-MEANS pode ser representado pela equação (10), na qual $X = \{x_i, i=1, \dots, n\}$ denota um conjunto de n pontos a serem agrupados em k *clusters*, definido por $C = \{c_k, k=1, \dots, k\}$ e u_k representa a média do *cluster* c_k (LIU; YU, 2009).

$$J(c_k) = \sum_{k=1}^k \sum_{x_i \in c_k} \|x_i - u_k\|^2. \quad (10)$$

3. MATERIAIS E MÉTODOS

Neste capítulo são descritos o método de pesquisa, a delimitação dos experimentos e os materiais utilizados, a forma como os experimentos foram conduzidos, bem como os parâmetros empregados no SE proposto e os critérios para a avaliação do seu desempenho.

3.1. Método de pesquisa

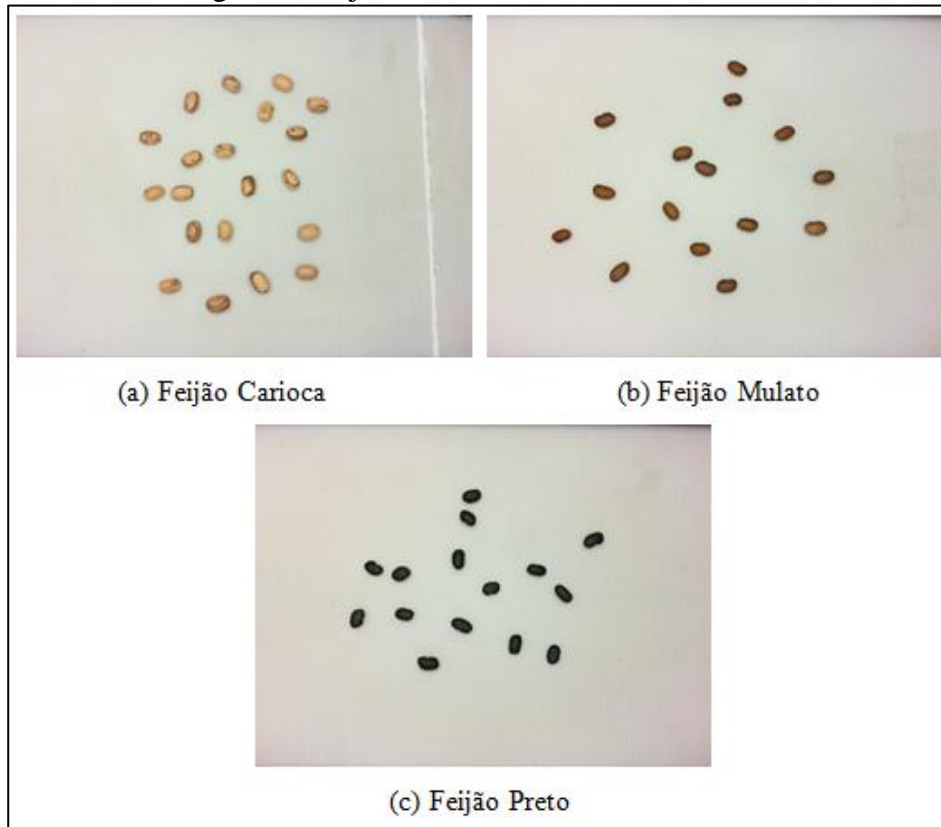
Este trabalho pode ser classificado, quanto a sua natureza, como uma pesquisa aplicada, pois tem por objetivo gerar conhecimentos específicos para a aplicação prática. Com relação à abordagem do problema, se enquadra como uma pesquisa quantitativa, uma vez que emprega técnicas estatísticas para expor, classificar e analisar informações (GIL, 1994).

3.2. Delimitação dos experimentos e materiais empregados

Os experimentos realizados tiveram como objetivo avaliar o desempenho do SE proposto no processo de classificação dos grãos de feijão. A classificação foi realizada com base na coloração da película de cada grão de feijão, de forma que fosse possível determinar a quantidade de misturas de grãos presentes em cada amostra de feijão carioca avaliada. Não pertenceu ao escopo dos experimentos, a classificação dos grãos de feijão a partir dos seus defeitos, ou ainda a aprovação ou reprovação de cada amostra, baseada nos limites de tolerância de misturas determinados para cada classe de feijão. A ideia é que somente após a avaliação de um conjunto de amostras, a informação da classe seja determinada.

Dessa forma, para a realização dos experimentos foram considerados os principais grãos de feijão comercializados no Brasil, pertencentes ao grupo I, das classes preto e cores. Mais especificamente, foram considerados os feijões carioca, mulato e preto (Figura 7).

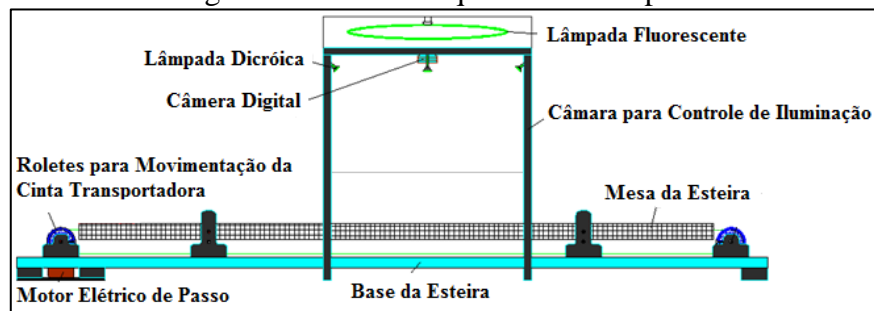
Figura 7: Feijões considerados neste trabalho.



Fonte: o autor.

Para a aquisição das amostras foi utilizado um aparato proposto por Pires (2012), ilustrado na Figura 8, composto por uma esteira com uma câmera digital acoplada, além de um esquema de iluminação composto por uma lâmpada circular fluorescente para gerar iluminação uniforme e mais quatro lâmpadas dicróicas para geração de iluminação local. Este aparato foi escolhido para a realização dos experimentos por apresentar características presentes em equipamentos encontrados em um processo industrial e por possuir um sistema de iluminação que possibilitasse a manipulação das sombras dos grãos de feijão, utilizadas pelo método proposto nesse trabalho.

Figura 8: Desenho esquemático do aparato.



Fonte: adaptado de Pires (2012).

Para implementação do SE foi utilizada a linguagem de programação C/C++, por se tratar de uma das linguagens mais populares e amplamente utilizada em processos de automação, dada a sua facilidade de integração com hardware. Também foram utilizadas as bibliotecas ProEikon, proposta por Kim (2011) e OpenCV, desenvolvida pela Intel (2000), destinadas ao Processamento de Imagens e Visão Computacional, além da rotina MGRANUL, proposta por Maruta (2011), empregada na detecção de grãos presentes em cada imagem digital.

3.3. Metodologia para a condução dos experimentos e parametrização do SE proposto

Nos experimentos com o SE proposto, foram realizados testes a partir de um conjunto composto por 100 imagens de amostras de feijão. Essas imagens foram divididas em 10 subconjuntos, a qual em cada subconjunto, quantidades específicas de feijões carioca, mulato e preto foram consideradas, havendo apenas a variação de suas posições em cada amostra. Todas as amostras foram compostas por 100 grãos de feijão cada uma.

Esses subconjuntos foram definidos de forma que as amostras contivessem diferentes misturas de grãos de feijão para avaliação. Na Tabela 2 é possível verificar os subconjuntos gerados para os experimentos.

Tabela 2: Subconjuntos de amostras gerados para a realização dos experimentos.

Subconjuntos	Total de imagens	Quantidade de grãos de feijão carioca	Quantidade de grãos de feijão preto	Quantidade de grãos de feijão mulato
1	10	100	0	0
2	10	95	0	5
3	10	95	5	0
4	10	90	5	5
5	10	85	5	10
6	10	85	10	5
7	10	80	10	10
8	10	85	0	15
9	10	85	15	0
10	10	70	15	15
Total	100	870	65	65

Fonte: o autor.

Os experimentos foram realizados considerando a classificação dos grãos de feijão tanto pelas RNA, detalhada no item 2.5.1, quanto pela técnica K-MEANS, detalhada no item 2.5.2.

Para a classificação a partir das RNA, foi considerada a arquitetura de rede *Multi-Layer Feedforward Network*, com uma camada oculta e aprendizado supervisionado. O treinamento da rede foi realizado a partir de um vetor de atributos, gerado com base em algumas amostras. O erro considerado para finalizar o treinamento da RNA foi de 0,0001.

Após o treinamento da rede, as imagens (amostras) de cada subconjunto formado para os experimentos foram submetidas para a classificação e cada grão de feijão detectado recebeu o rótulo ‘C’ quando classificado como carioca, ‘M’ quando classificado como mulato e ‘P’ quando classificado como preto.

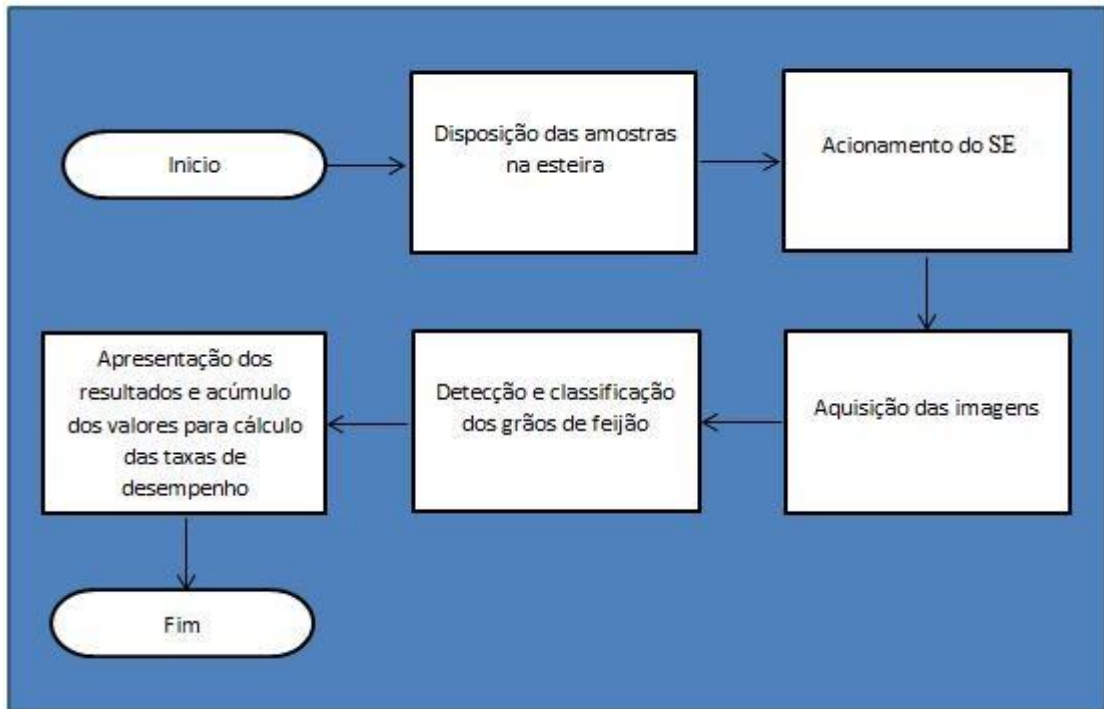
Para a classificação a partir da técnica K-MEANS, foram definidos três centróides com base no vetor de atributos, gerado a partir de algumas amostras. Neste caso, cada centróide foi criado para representar cada classe de feijão considerada nos experimentos. As amostras de cada subconjunto formado para os experimentos também foram submetidas para a classificação, como ocorrido no experimento usando RNA.

Foram realizados experimentos com as imagens no sistema de cores RGB e também no sistema de cores CIELAB. Para o sistema de cores CIELAB, o valor de L^* foi descartado para verificar se a subtração do parâmetro de luminosidade traria alguma vantagem com relação ao sistema de cores RGB.

Na condução dos experimentos, o seguinte fluxo de atividades (ilustrado na Figura 9) foi adotado:

- Cada amostra de feijão é disposta de forma manual na esteira;
- O SE é acionado, fazendo com que a esteira conduza a amostra até a área de captura das imagens;
- Cinco imagens da amostra são capturadas automaticamente pela câmera digital;
- O SE processa a imagem detectando e classificando os grãos de feijão presentes na amostra;
- Os resultados da classificação são apresentados por meio da interface gráfica desenvolvida e são acumulados para o cálculo das taxas de acertos, precisão e acurácia;

Figura 9: Fluxograma dos experimentos realizados.



Fonte: o autor.

Para o funcionamento do SE proposto, alguns parâmetros precisaram ser ajustados. Na Tabela 3 estão descritos os valores assumidos para esses parâmetros, os quais foram obtidos experimentalmente.

Tabela 3: Valores assumidos para os parâmetros durante os experimentos.

Parâmetro	Descrição	Valor Adotado
l_{borda}	Limiar para a intensidade dos níveis de cinza das bordas.	0,4
$l_{composicao}$	Limiar para a intensidade dos níveis de cinza da imagem com iluminação uniforme.	0,2
l_{corr}	Limiar para a taxa de correlação.	0,15
l_{sobrep}	Limiar para o valor de sobreposição	0,1
r_{ij}	Raios das elipses	$r_{11} = 38 \mid r_{12} = 27$ $r_{21} = 35 \mid r_{22} = 25$ $r_{31} = 20 \mid r_{32} = 30$
s_i	Escala das elipses	$s_1 = 0,9$ $s_2 = 1,0$ $s_3 = 1,1$

Fonte: o autor.

3.4. Avaliação de desempenho

A avaliação de desempenho do SE foi realizada a partir das taxas de acertos, precisão e acurácia, calculadas com base na matriz de confusão ou tabela de contingência proposta por FAWCETT (2006). A partir desta matriz de confusão, quatro possíveis categorias podem ser atribuídas para cada classificação realizada: Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN), em que:

- VP refere-se à situação em que uma instância do objeto procurado existe na imagem e é detectada pelo algoritmo. Para o SE, o VP foi atribuído para os casos em que o grão de feijão foi detectado e classificado corretamente.
- VN refere-se à situação em que não existe uma instância do objeto procurado na imagem e o algoritmo não a detecta. Neste trabalho, como todas as instâncias de feijão presentes nas imagens deveriam ser classificadas, esta categoria recebeu o valor 0 (zero) para o cálculo da taxa de acurácia, descrita adiante.
- FP refere-se à situação em que não existe uma instância do objeto procurado na imagem, porém é detectada indevidamente pelo algoritmo. Para os feijões, o FP foi atribuído para os casos em que o grão foi detectado, porém classificado incorretamente, ou ainda quando houve alguma detecção indevida.
- FN refere-se à situação em que uma instância do objeto procurado existe na imagem, porém não é detectada pelo algoritmo. Neste trabalho, o FN foi atribuído para os casos em que o grão de feijão deixou de ser detectado e classificado corretamente.

Após computar os valores VP, VN, FP e FN com base na amostra de grãos de feijão analisada, as taxas de acertos, precisão e acurácia podem ser calculadas a partir das equações (11), (12) e (13):

$$\text{Taxa de Acertos} = \frac{VP}{VP+FN} \times 100, \quad (11)$$

$$\text{Taxa de Precisão} = \frac{VP}{VP+FP} \times 100, \quad (12)$$

$$\text{Taxa de Acurácia} = \frac{VP+VN}{VP+VN+FP+FN} \times 100 . \quad (13)$$

Na qual a taxa de acertos corresponde à razão entre o número de classificações realizadas corretamente e o total de instâncias do objeto presentes na imagem. A taxa de precisão representa a razão entre o número de classificações realizadas corretamente e o total de instâncias classificadas pelo algoritmo e a taxa de acurácia representa a fração dos casos corretamente classificados, ou seja, a proximidade do total de classificações realizadas corretamente com relação ao total de classificações que deveriam ser realizadas (ARAÚJO, 2009).

4. RESULTADOS E DISCUSSÕES

Neste capítulo é descrito o funcionamento do SE proposto, bem como os resultados obtidos nos experimentos realizados.

4.1. Detalhamento do sistema especialista proposto.

O SE proposto (Figura 10) é composto por diferentes etapas para classificar cada grão de feijão presente em uma amostra. Trata-se do conjunto de etapas típicas em um SRPV, ou seja, aquisição de imagens, pré-processamento, segmentação, representação e descrição e reconhecimento e interpretação, as quais são detalhadas nas subseções a seguir.

Figura 10: Aparato para aquisição das imagens e a interface gráfica para utilização do SE.



Fonte: o autor.

4.1.1. Aquisição de imagens

O primeiro passo é realizar a aquisição das imagens da amostra submetida para análise. Para isso, o SE foi integrado ao aparato proposto por Pires (2012), de modo que o sistema detecta automaticamente a presença da amostra disposta na esteira e realiza a aquisição das imagens.

Para a detecção automática da amostra a ser analisada, o sistema avalia a variação de gradiente ao longo dos quadros capturados pela câmera. Desta forma, é possível identificar o ponto inicial em que a amostra atinge a área focal da câmera e o momento final.

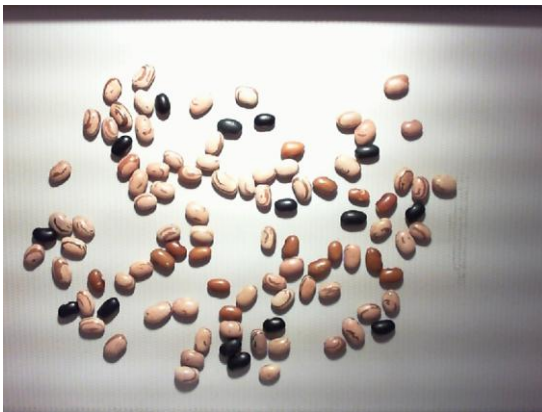
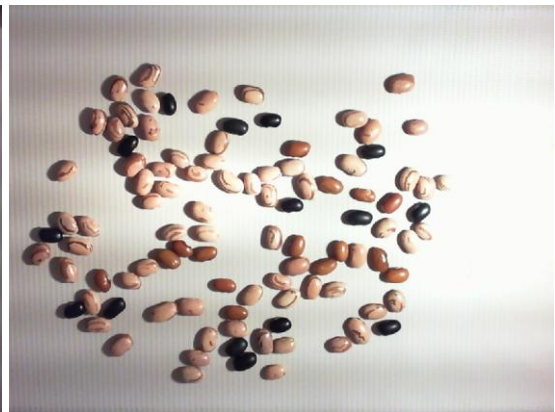
Ao identificar a amostra, o movimento da esteira é temporariamente interrompido e cinco imagens coloridas são capturadas da amostra. A primeira aquisição é realizada considerando iluminação uniforme.

A iluminação uniforme elimina a presença de sombras entre os grãos de feijão e é gerada a partir da lâmpada circular fluorescente, localizada na parte superior da câmara de iluminação. As demais aquisições ocorrem a partir das quatro lâmpadas dicróicas. Para cada imagem capturada, uma única lâmpada dicróica é acesa, gerando propositalmente sombra entre os grãos de feijão.

Na maior parte dos SRPV, as sombras são indesejadas por influenciar na detecção e análise dos objetos. Por este motivo, é comum que somente a iluminação uniforme seja utilizada.

Neste trabalho, entretanto, as sombras são geradas propositalmente para que as bordas de cada grão de feijão sejam destacadas, auxiliando na sua detecção, mesmo quando estão próximos ou encostados uns aos outros. A Figura 11 demonstra um exemplo de cinco imagens capturadas a partir de uma mesma amostra. Essas imagens são denotadas por I_0 , I_1 , I_2 , I_3 e I_4 .

Figura 11: Cinco imagens com variação na iluminação, capturadas de uma mesma amostra.

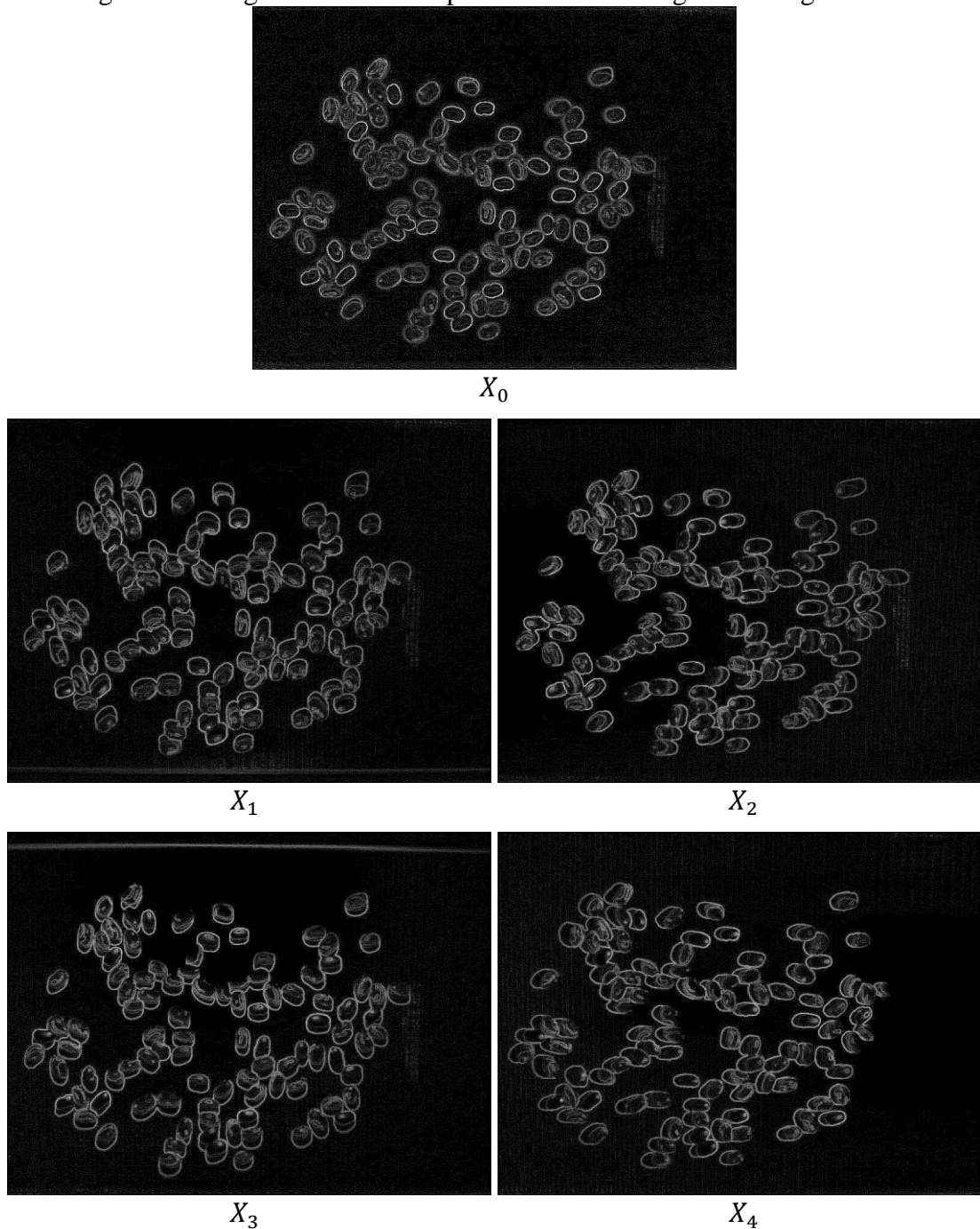
 I_0  I_1  I_2  I_3  I_4

Fonte: o autor.

4.1.2. Pré-processamento

Esta etapa se inicia convertendo cada imagem colorida I_i em sua correspondente em níveis de cinzas G_i . Na sequência, é calculado o gradiente de cada imagem G_i , por meio da equação 4, visando detectar as bordas de cada grão. Cinco novas imagens são geradas (Figura 12), denotadas por X_0 , X_1 , X_2 , X_3 e X_4 .

Figura 12: Imagens da amostra após o cálculo da magnitude do gradiente.

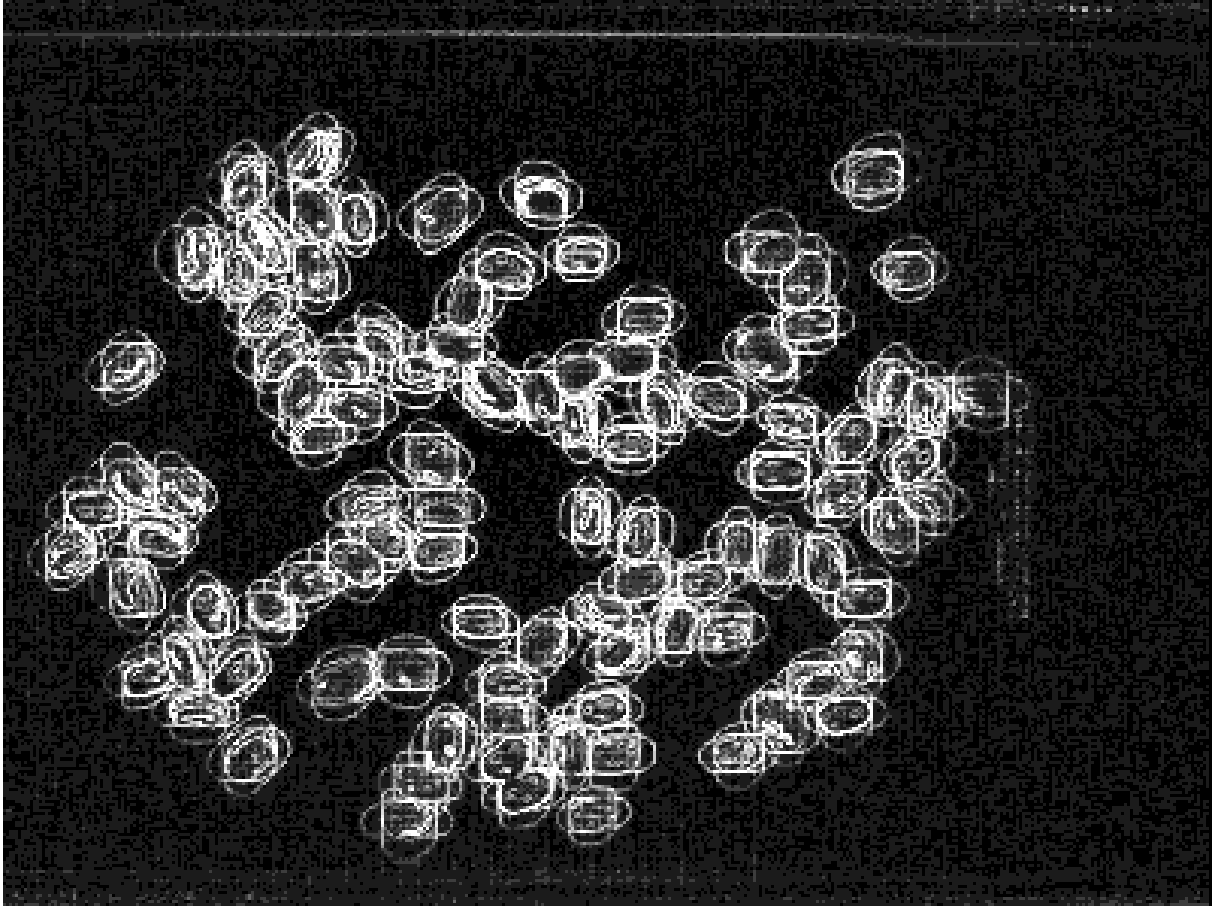


Fonte: o autor.

Em seguida, a máxima magnitude do gradiente é calculada a partir das imagens X_i (equação 15), gerando uma nova imagem, denominada Z , conforme Figura 13.

$$Z(x, y) = \text{MAX}_{i=1}^5 [X_i(x, y)]. \quad (15)$$

Figura 13: Imagem Z .

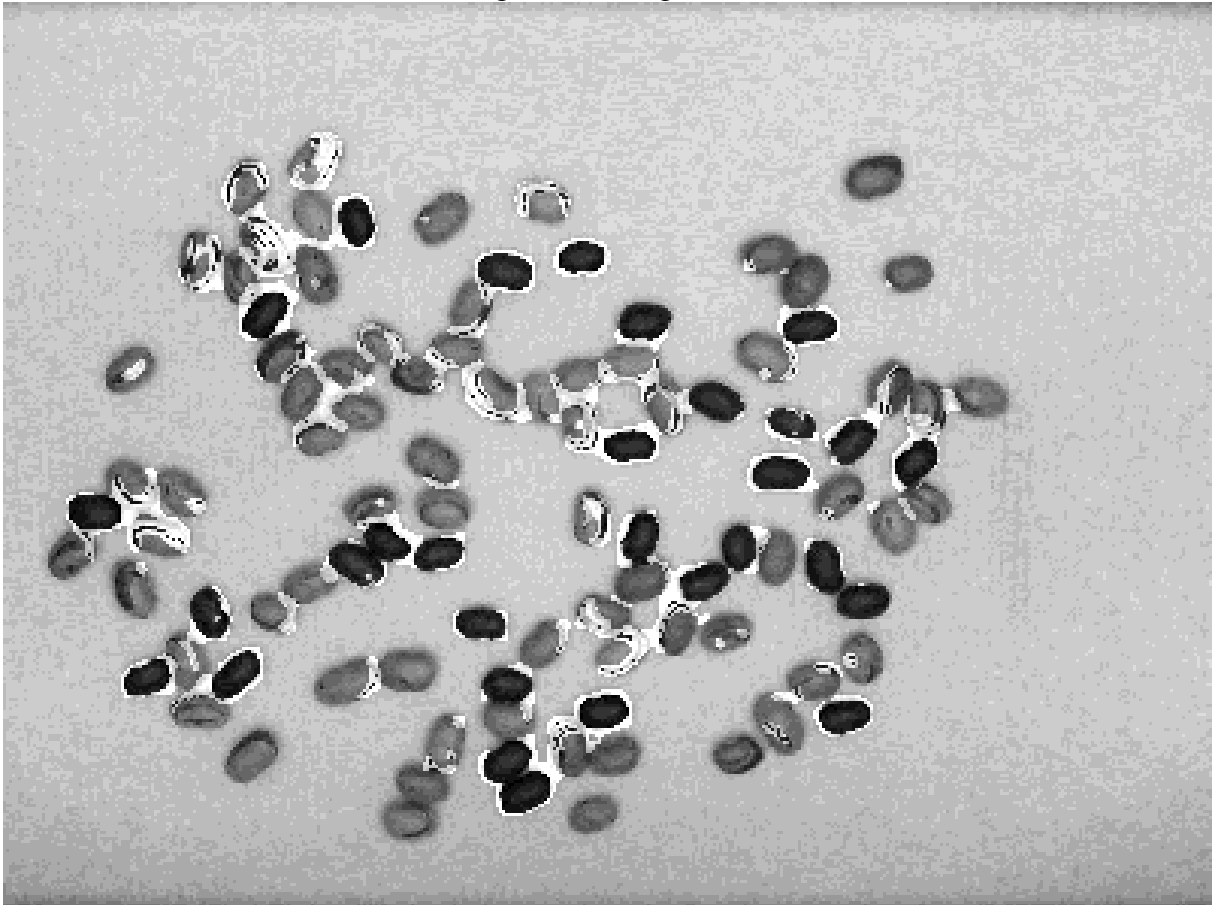


Fonte: o autor.

Na sequência, um filtro gaussiano é aplicado na imagem Z , com o objetivo de reduzir ruídos e unificar as bordas próximas a partir do efeito de suavização. Uma nova imagem, denominada S , é gerada a partir da composição dos pixels da imagem Z , que representam as bordas com maior intensidade de níveis de cinza, com os pixels da imagem com iluminação uniforme G_0 . Para este processo, dois parâmetros são considerados, o l_{borda} , que representa o limiar para determinar as bordas que possuem alta intensidade de níveis de cinza na imagem Z , e o parâmetro $l_{composição}$, que representa o limiar para determinar quais pixels da imagem G_0 devem compor a nova imagem S .

Estes dois parâmetros podem variar de 0 a 1 e para os experimentos realizados foram assumidos os valores apresentados na Tabela 3. O resultado deste processo pode ser observado na Figura 14, em que os pixels onde grãos de feijão estão próximos ou encostados ficam destacados.

Figura 14: Imagem S .



Fonte: o autor.

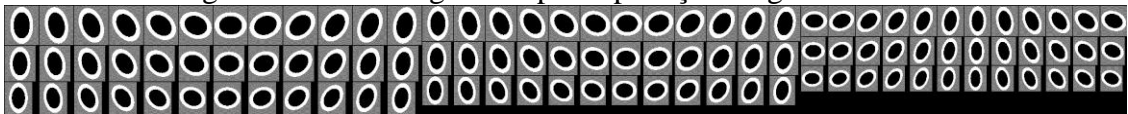
4.1.3. Segmentação/detecção dos grãos

Nesta etapa, a técnica de granulometria (equações 6 a 9) é aplicada na imagem S , gerando como saída o coeficiente da correlação cruzada entre os elementos presentes na imagem analisada e os núcleos gerados para a comparação, além das coordenadas espaciais do centro de cada grão detectado. Para a geração dos núcleos, foram considerados três tamanhos distintos de elipses (e_1 , e_2 e e_3), determinados de acordo com o tamanho médio dos grãos de feijão analisados. Os pares de valores assumidos para os raios r_{ij} de cada elipse e_i estão descritos na Tabela 3.

Para que a correlação cruzada entre as elipses e_i e os grãos de feijão da imagem S fosse invariante à rotação, foram consideradas doze rotações, com variação de 15 graus entre cada uma, contemplando um intervalo de 0 a 180 graus. Também foram consideradas três escalas (s_1 , s_2 e s_3), para que a comparação fosse invariante à escala.

Ao todo, cento e oito núcleos foram gerados para esta etapa de detecção dos grãos, conforme ilustrado na Figura 15.

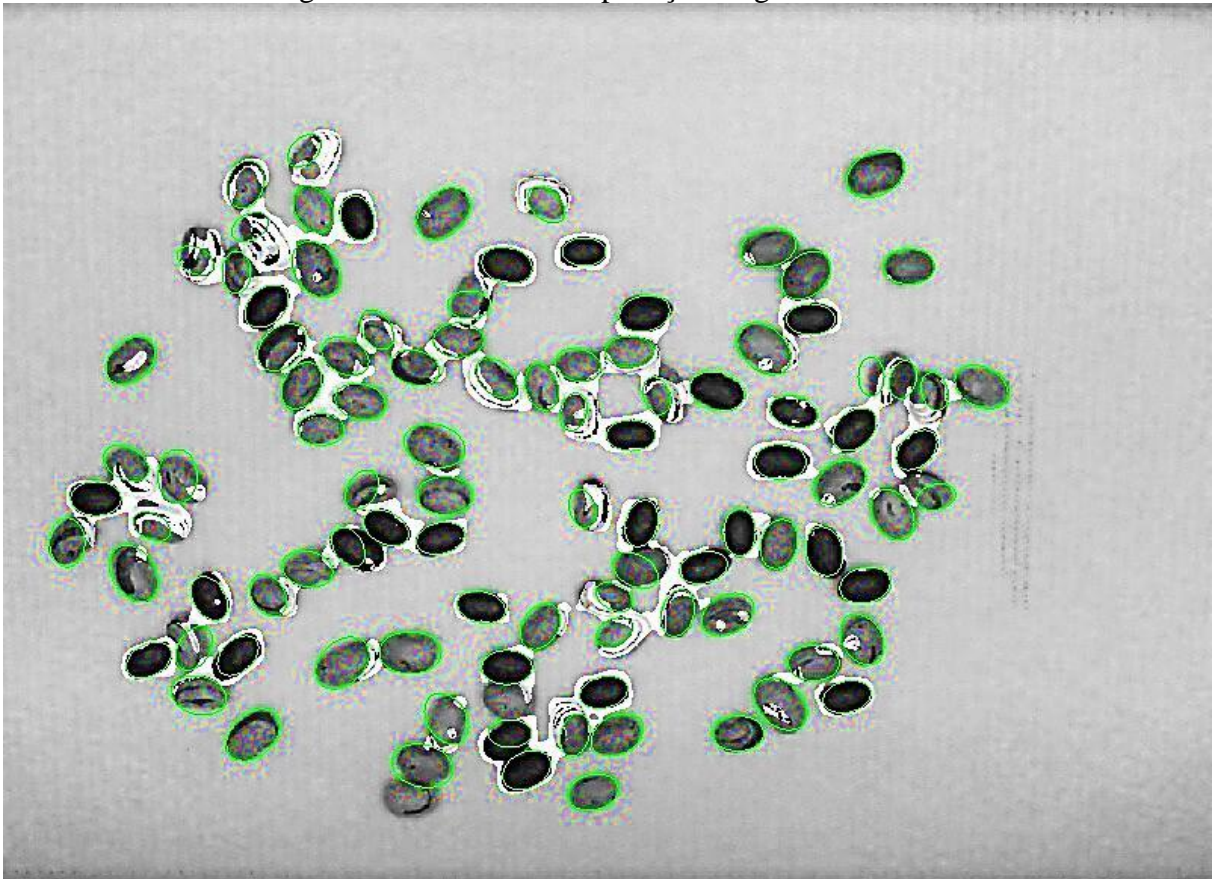
Figura 15: Núcleos gerados para aplicação da granulometria.



Fonte: o autor.

Na sequência, são descartados todos os elementos que apresentam baixa correlação ou que apresentam alguma sobreposição com outros elementos com maior correlação, utilizando os parâmetros l_{corr} e l_{sobrep} . A Figura 16 apresenta o resultado final da aplicação do processo de granulometria, onde todos os grãos de feijão detectados estão contornados por uma linha na cor verde.

Figura 16: Resultado da aplicação da granulometria.



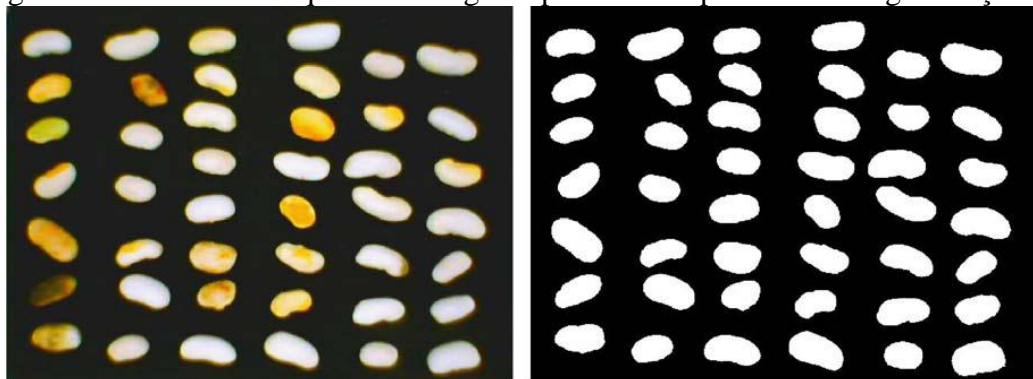
Fonte: o autor.

A etapa de pré-processamento aplicada na geração da imagem S , a qual envolve o esquema de iluminação com as lâmpadas dicróicas para geração das sombras em diferentes direções durante a aquisição das imagens I_1 a I_4 , é fundamental para que se obtenha um bom desempenho na detecção dos grãos através da aplicação da técnica de granulometria.

Não obstante, a abordagem para detecção de grãos proposta neste trabalho permite contornar uma limitação existente nos trabalhos encontrados na literatura para análise e classificação de sementes e grãos, que é a necessidade de distanciamento dos objetos (sementes/grãos) para facilitar o processo de segmentação/detecção. Cabe ressaltar que essa limitação dificulta a aplicabilidade de tais sistemas em processos industriais, tendo em vista a necessidade de um equipamento adicional para fazer automaticamente a disposição dos grãos/sementes em cada amostra a ser analisada.

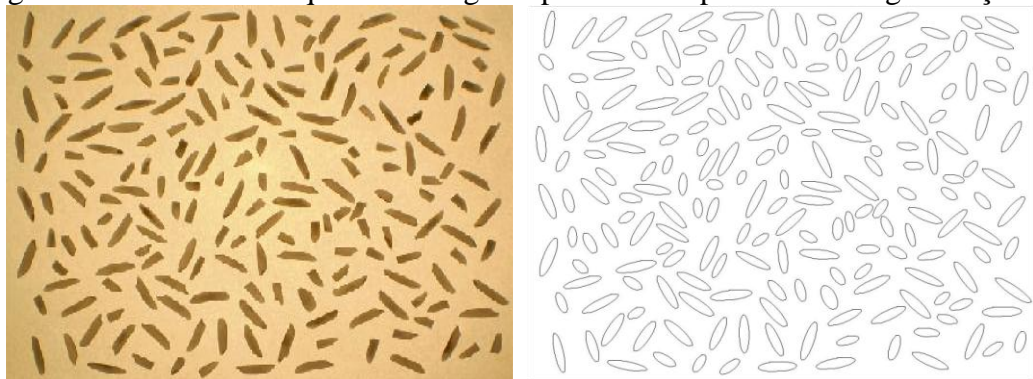
Nos trabalhos de Kiliç et al. (2007) e Aggarwal e Mohan (2010), por exemplo, é possível verificar o distanciamento dos grãos, conforme ilustram as Figuras 17 e 18.

Figura 17: Disposição dos grãos de feijão em uma amostra analisada. À direita a imagem original da amostra e à esquerda a imagem após sofrer o processo de segmentação.



Fonte: Kiliç et al. (2007).

Figura 18: Disposição dos grãos de arroz em uma amostra analisada. À direita a imagem original da amostra e à esquerda a imagem após sofrer o processo de segmentação.



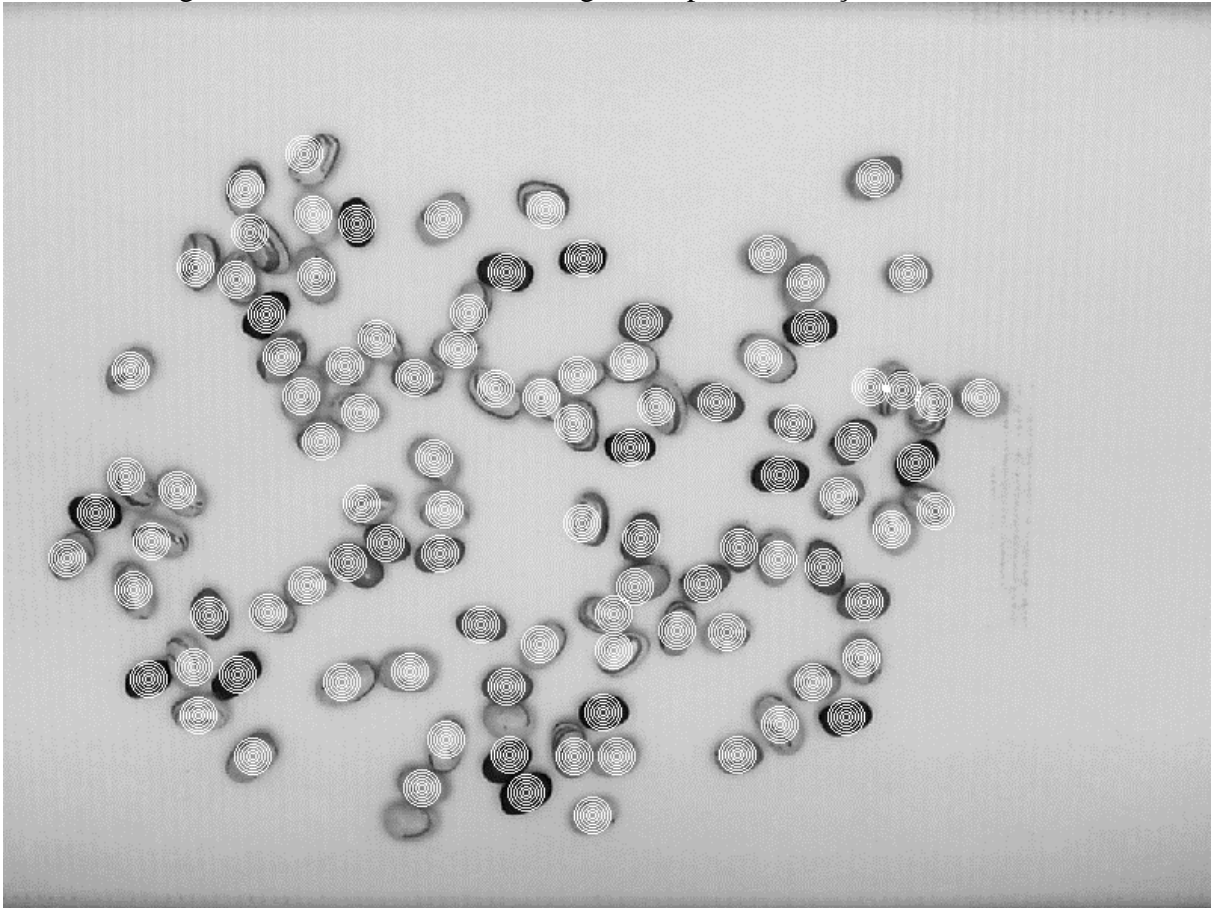
Fonte: Aggarwal e Mohan, (2010).

4.1.4. Representação e descrição

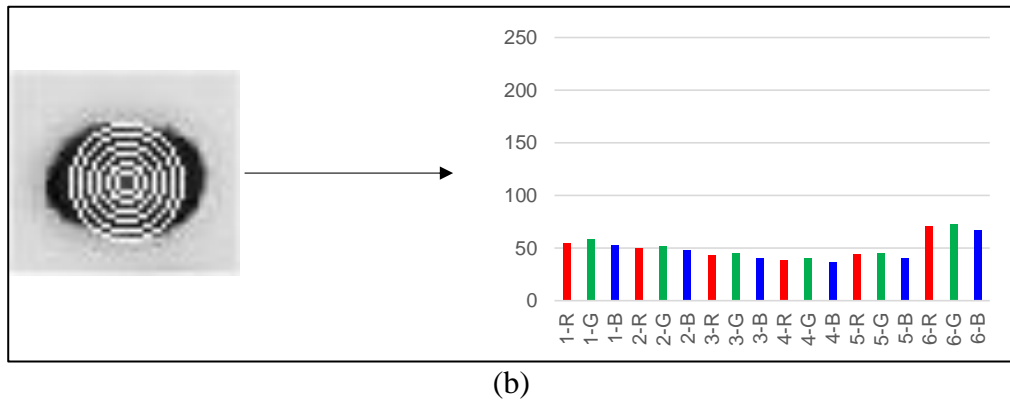
Nesta etapa realiza-se a extração dos atributos que descrevem cada grão de feijão detectado. Para isso, são gerados círculos concêntricos a partir das coordenadas espaciais do centro de cada grão, obtidas na etapa anterior. Neste caso, seis círculos são criados com distância de um pixel entre cada um.

Em seguida, a média das componentes de cor (RGB ou Lab) ao longo de cada círculo é armazenada em um vetor de atributos. A Figura 19(a) mostra os círculos concêntricos gerados para cada grão de feijão enquanto a Figura 19(b) ilustra o histograma gerado a partir das médias de R, G e B, extraídas das projeções circulares feitas em um único grão.

Figura 19: Círculos concêntricos gerados para a extração dos atributos.



(a)



Fonte: o autor.

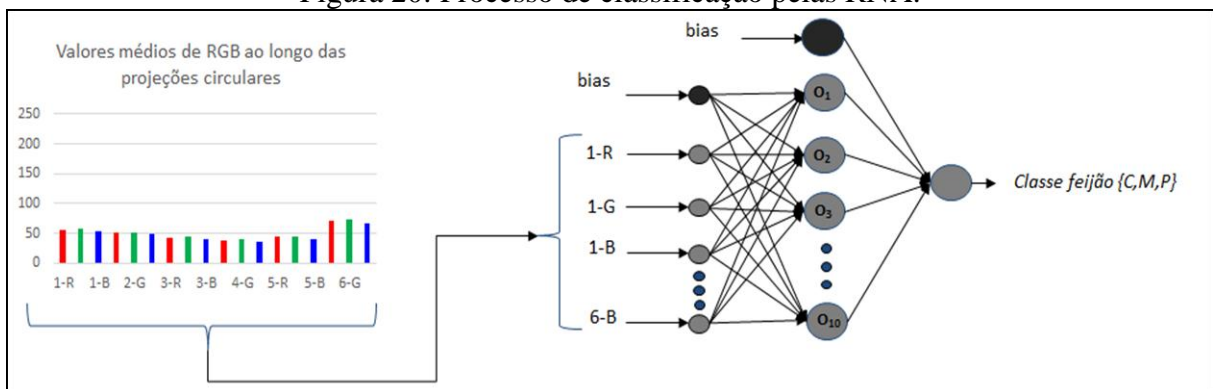
No caso da extração de atributos usando o sistema de cores Lab, o processo é análogo, ou seja, o vetor de atributos é computado a partir das médias das componentes L, a e b ao longo das projeções circulares em cada grão.

4.1.5. Reconhecimento e interpretação

Nesta última etapa, cada grão de feijão é classificado com base no vetor de atributos gerado no passo anterior. O SE foi desenvolvido de forma que a classificação possa ser realizada tanto pelas RNA, quanto pela técnica K-MEANS, com o objetivo de verificar qual técnica de classificação apresenta os melhores resultados.

Para a classificação através das RNA, as médias das componentes de cor (RGB ou Lab) de cada grão de feijão são submetidas para a rede neural, de modo que os pesos da camada oculta sejam aplicados, gerando como resultado, a classe a qual o grão de feijão pertence. A Figura 20 ilustra este processo.

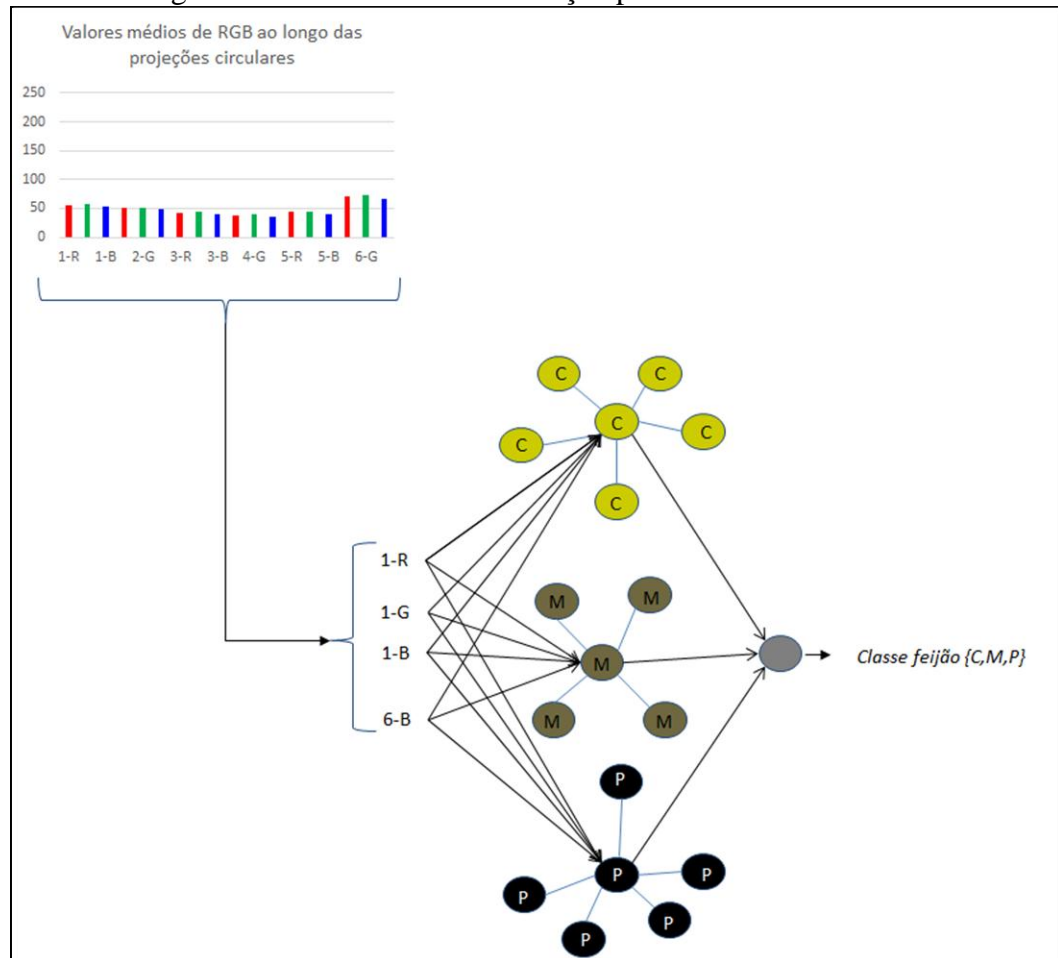
Figura 20: Processo de classificação pelas RNA.



Fonte: o autor.

Para a classificação através da técnica K-MEANS, é calculado o somatório das diferenças entre as médias das componentes de cor (RGB ou Lab) de cada grão de feijão, com relação a cada centróide gerado na etapa de treinamento. Neste caso, a classe do grão de feijão é determinada a partir do centróide que resulta na menor diferença global. A Figura 21 exemplifica este processo.

Figura 21: Processo de classificação pela técnica K-MEANS.



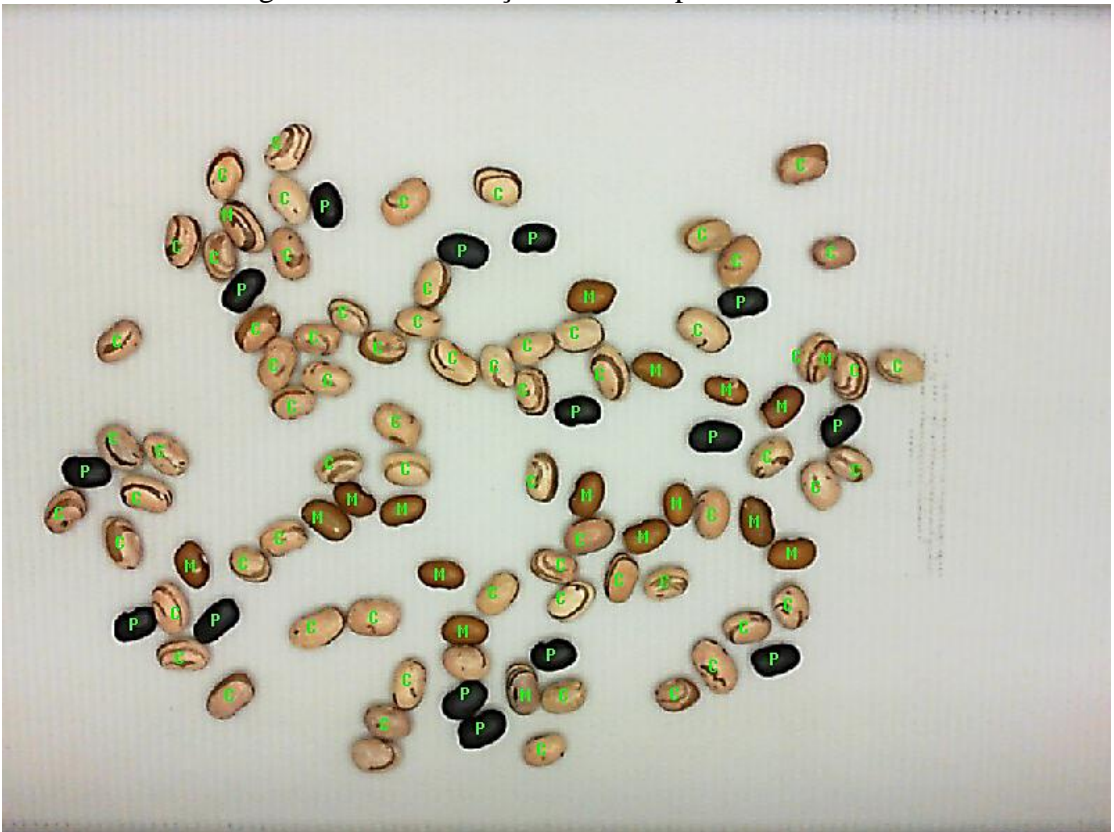
O resultado da aplicação dos classificadores RNA e K-MEANS em amostras de feijão pode ser observado a partir da Figura 22, que ilustra a classificação para uma amostra usando RNA, enquanto que a Figura 23 apresenta a classificação para a mesma amostra a partir da técnica K-MEANS.

Figura 22: Classificação realizada pela RNA.



Fonte: o autor.

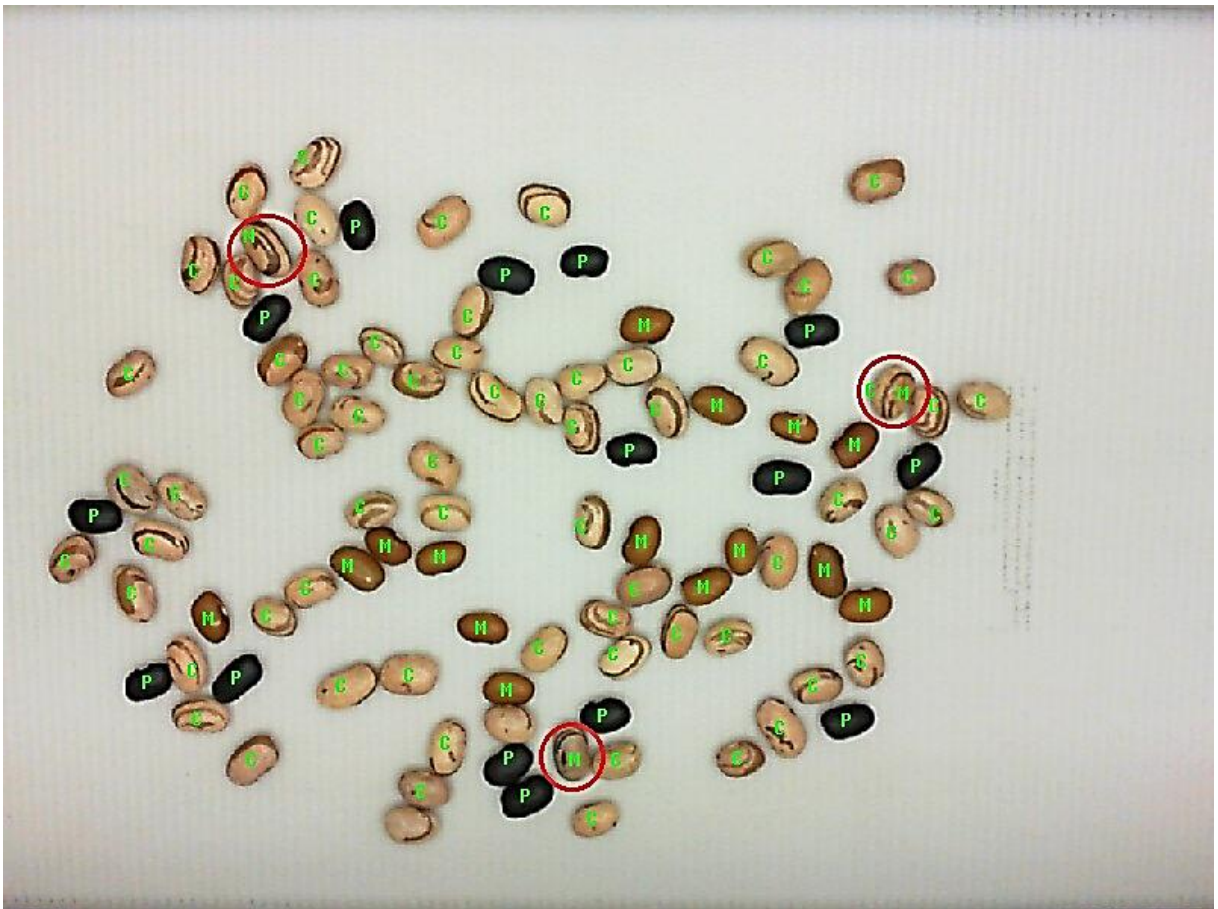
Figura 23: Classificação realizada por K-MEANS.



Fonte: o autor.

Para a amostra utilizada como exemplo, ocorreram três diferenças entre as classificações realizadas por RNA e K-MEANS. Estas diferenças estão destacadas na Figura 24 e referem-se a erros gerados pela técnica K-MEANS, que classificou incorretamente três grãos de feijão carioca como mulato. Também é possível observar que o grão de feijão destacado mais à direita na imagem foi detectado duas vezes e, portanto, classificado como se fossem dois grãos. Contudo, a RNA classificou este grão como carioca, enquanto que K-MEANS produziu duas diferentes classificações, carioca e mulato.

Figura 24: Diferenças geradas na classificação a partir das técnicas RNA e K-MEANS, para uma mesma amostra.



Fonte: o autor.

4.2. Experimentos realizados com o SE proposto

Ao todo, quatro tipos de experimentos foram realizados considerando os seguintes cenários: classificação K-MEANS e sistema de cores RGB; classificação K-MEANS e sistema de cores CIELAB; classificação com RNA e sistema de cores RGB; e classificação com RNA e sistema de cores CIELAB.

A Tabela 4 mostra os resultados obtidos com os experimentos realizados utilizando a classificação K-MEANS e imagens no sistema de cores RGB.

Tabela 4: Resultados experimentais obtidos com classificação K-MEANS e sistemas de cores RGB.

Subconjunto	Quantidade de grãos de feijão carioca em cada amostra	Quantidade de grãos de feijão preto em cada amostra	Quantidade de grãos de feijão mulato em cada amostra	VP	FP	FN	VN	Taxa de acertos (%)	Precisão (%)	Acurácia (%)
1	100	0	0	945	47	55	0	94,50	95,26	90,26
2	95	0	5	959	56	41	0	95,90	94,48	90,81
3	95	5	0	940	57	60	0	94,00	94,28	88,93
4	90	5	5	963	6	37	0	96,30	99,38	95,73
5	85	5	10	962	50	38	0	96,20	95,06	91,62
6	85	10	5	983	29	17	0	98,30	97,13	95,53
7	80	10	10	992	20	8	0	99,20	98,02	97,25
8	85	0	15	953	40	47	0	95,30	95,97	91,63
9	85	15	0	931	49	69	0	93,10	95,00	88,75
10	70	15	15	956	40	44	0	95,60	95,98	91,92
Total	870	65	65	9584	394	416	0	95,84	96,05	92,21

Fonte: o autor.

Nestes experimentos foi possível observar que as três taxas avaliadas tiveram um resultado médio acima de 90%. Os melhores resultados foram obtidos com o subconjunto 7, para o qual, as taxas de acertos, precisão e acurácia foram de 99,20%, 98,02% e 97,25%, respectivamente. Os piores resultados foram obtidos com o subconjunto 9, em que a taxa de acertos foi de 93,10%, a taxa de precisão foi de 95% e a taxa de acurácia foi de 88,75%. Ao analisar as amostras do subconjunto 9, percebeu-se que apesar da iluminação ser controlada no momento da aquisição das imagens, houve uma falha de sincronia entre o desligamento de uma lâmpada e o acendimento de outra, no momento da aquisição da imagem com iluminação uniforme, o que resultou em uma grande área de sombra na parte inferior da imagem da

amostra 5, conforme apresentado na Figura 25. Essa área de sombra não esperada resultou em 50,85% de todos os erros obtidos neste subconjunto.

Figura 25: Imagem com iluminação uniforme pertencente à amostra 5 do subconjunto 9.



Fonte: o autor.

Com base nos experimentos realizados, observa-se que a utilização da técnica de classificação K-MEANS com imagens no sistema de cores RGB apresentou bons resultados, com valores médios acima de 90% para as três taxas avaliadas.

Na Tabela 5 são apresentados os resultados obtidos com os experimentos realizados utilizando a classificação K-MEANS e imagens no sistema de cores CIELAB.

Tabela 5: Resultados experimentais obtidos com classificação K-MEANS e sistemas de cores CIELAB.

Subconjuntos	Quantidade de grãos de feijão carioca em cada amostra	Quantidade de grãos de feijão preto em cada amostra	Quantidade de grãos de feijão mulato em cada amostra	VP	FP	FN	VN	Taxa de acertos (%)	Precisão (%)	Acurácia (%)
1	100	0	0	935	57	65	0	93,50	94,25	88,46
2	95	0	5	940	75	60	0	94,00	92,61	87,44
3	95	5	0	939	58	61	0	93,90	94,18	88,75
4	90	5	5	953	50	47	0	95,30	95,01	90,76
5	85	5	10	947	65	53	0	94,70	93,58	88,92
6	85	10	5	940	72	60	0	94,00	92,89	87,69
7	80	10	10	941	71	59	0	94,10	92,98	87,86
8	85	0	15	943	50	57	0	94,30	94,96	89,81
9	85	15	0	916	64	84	0	91,60	93,47	86,09
10	70	15	15	941	55	59	0	94,10	94,48	89,19
Total	870	65	65	9395	617	605	0	93,95	93,84	88,49

Fonte: o autor.

Nestes experimentos foi possível observar que as três taxas avaliadas tiveram um resultado médio inferior ao resultado obtido com a classificação K-MEANS com sistema de cores RGB, principalmente para taxa de acurácia, que ficou abaixo de 90%. Os melhores resultados foram obtidos com o subconjunto 4, para o qual, as taxas de acertos, precisão e acurácia foram de 95,30%, 95,01% e 90,76%, respectivamente. Os piores resultados foram obtidos com o subconjunto 9, em que a taxa de acertos foi de 91,60%, a taxa de precisão foi de 93,47% e a taxa de acurácia foi de 86,09%. Como foram utilizadas as mesmas amostras para todos os experimentos, a imagem com iluminação uniforme da amostra 5 também foi o motivo pelo qual o subconjunto 9 tivesse o pior resultado. Neste caso, este problema representou 21,62% de todos os erros obtidos neste subconjunto.

Independente deste fator, o resultado médio obtido possibilitou observar que o emprego do sistema de cores CIELAB não proporcionou vantagens com relação ao sistema de cores RGB no processo de classificação pela técnica K-MEANS.

Na Tabela 6 são demonstrados os resultados obtidos com os experimentos realizados utilizando RNA para a classificação e imagens no sistema de cores RGB.

Tabela 6: Resultados experimentais obtidos com classificação através da RNA e sistemas de cores RGB.

Subconjuntos	Quantidade de grãos de feijão carioca em cada amostra	Quantidade de grãos de feijão preto em cada amostra	Quantidade de grãos de feijão mulato em cada amostra	VP	FP	FN	VN	Taxa de acertos (%)	Precisão (%)	Acurácia (%)
1	100	0	0	983	9	17	0	98,30	99,09	97,42
2	95	0	5	987	28	13	0	98,70	97,24	96,01
3	95	5	0	970	27	30	0	97,00	97,29	94,45
4	90	5	5	991	12	9	0	99,10	98,80	97,92
5	85	5	10	990	22	10	0	99,00	97,83	96,87
6	85	10	5	998	14	2	0	99,80	98,62	98,42
7	80	10	10	997	15	3	0	99,70	98,52	98,23
8	85	0	15	990	3	10	0	99,00	99,70	98,70
9	85	15	0	971	9	29	0	97,10	99,08	96,23
10	70	15	15	985	11	15	0	98,50	98,90	97,43
Total	870	65	65	9862	150	138	0	98,62	98,50	97,16

Fonte: o autor.

Nestes experimentos foi possível observar que as três taxas avaliadas tiveram um resultado médio acima de 95%. Os melhores resultados foram obtidos com o subconjunto 8, para o qual, as taxas de acertos, precisão e acurácia foram de 99%, 99,70% e 98,70% respectivamente. Os piores resultados foram obtidos com o subconjunto 3, em que a taxa de acertos foi de 97%, a taxa de precisão foi de 97,29% e a taxa de acurácia foi de 94,45%. Ao analisar as imagens do subconjunto 3, foi verificado que a amostra 1 também sofreu um problema de sincronismo entre a troca de iluminação e aquisição da imagem, da qual gerou uma área de sombra na parte inferior da imagem com iluminação uniforme, o que resultou em 63,16% de todos os erros obtidos neste subconjunto.

Na figura 26 é possível observar a faixa de sombra não esperada, presente na parte inferior da imagem com iluminação uniforme pertencente à amostra 1 do subconjunto 3.

Figura 26: Imagem com iluminação uniforme pertencente à amostra 01 do subconjunto 3.



Fonte: o autor.

Observa-se a partir dos experimentos realizados, que a utilização da RNA para a classificação com imagens no sistema de cores RGB gerou os melhores resultados. Neste caso, tanto os resultados individuais de cada subconjunto, quanto os resultados médios para cada taxa avaliada foram acima de 95%, com exceção apenas para a taxa de acurácia do subconjunto 3, influenciada pelo problema de sincronismo entre iluminação e o processo de aquisição das imagens.

Na Tabela 7 são demonstrados os resultados obtidos com os experimentos realizados utilizando RNA para a classificação e imagens no sistema de cores CIELAB.

Tabela 7: Resultados experimentais obtidos com classificação através da RNA e sistemas de cores CIELAB.

Subconjunto	Quantidade de grãos de feijão carioca em cada amostra	Quantidade de grãos de feijão preto em cada amostra	Quantidade de grãos de feijão mulato em cada amostra	VP	FP	FN	VN	Taxa de acertos (%)	Precisão (%)	Acurácia (%)
1	100	0	0	982	10	18	0	98,20	98,99	97,23
2	95	0	5	983	32	17	0	98,30	96,85	95,25
3	95	5	0	979	18	21	0	97,90	98,19	96,17
4	90	5	5	982	21	18	0	98,20	97,91	96,18
5	85	5	10	963	49	37	0	96,30	95,16	91,80
6	85	10	5	990	22	10	0	99,00	97,83	96,87
7	80	10	10	968	44	32	0	96,80	95,65	92,72
8	85	0	15	940	53	60	0	94,00	94,66	89,27
9	85	15	0	972	8	28	0	97,20	99,18	96,43
10	70	15	15	947	49	53	0	94,70	95,08	90,28
Total	870	65	65	9706	306	294	0	97,06	96,94	94,18

Fonte: o autor.

Nestes experimentos as três taxas avaliadas tiveram um resultado médio acima de 90%. Os melhores resultados foram obtidos com o subconjunto 1, para o qual, as taxas de acertos, precisão e acurácia foram de 98,20%, 98,99% e 97,23% respectivamente. Os piores resultados foram obtidos com o subconjunto 8, em que a taxa de acertos foi de 94%, a taxa de precisão foi de 94,66% e a taxa de acurácia foi de 89,27%. Neste caso, verificou-se que o subconjunto 8 apresentou os piores resultados devido a uma alta quantidade de FP associados à classe carioca e uma quantidade também alta de FN associados à classificação dos grãos de feijão mulato. Com isso, foi possível observar que o emprego do sistema de cores CIELAB também não proporcionou vantagens com relação ao sistema de cores RGB no processo de classificação pela técnica de RNA.

A Tabela 8 sumariza os resultados obtidos nos experimentos. Como pode ser visto, no geral, a classificação com RNA apresentou melhores resultados do que a classificação por K-MEANS.

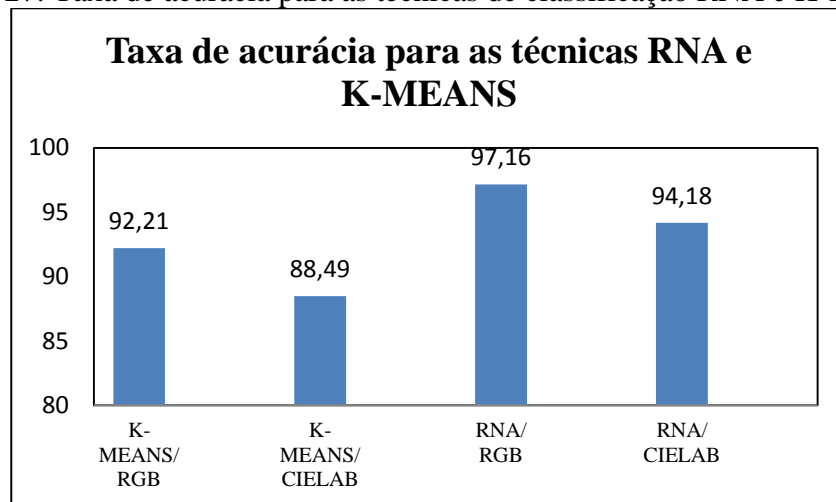
Tabela 8: Resultados experimentais sumarizados.

Tipo de Experimento	Quantidade de grãos de feijão carioca em cada amostra	Quantidade de grãos de feijão preto em cada amostra	Quantidade de grãos de feijão mulato em cada amostra	VP	FP	FN	VN	Taxa de acertos (%)	Precisão (%)	Acurácia (%)
K-MEANS/RGB	870	65	65	9584	394	416	0	95,84	96,05	92,21
K-MEANS/CIELAB	870	65	65	9395	617	605	0	93,95	93,84	88,49
RNA/RGB	870	65	65	9862	150	138	0	98,62	98,50	97,16
RNA/CIELAB	870	65	65	9706	306	294	0	97,06	96,94	94,18

Fonte: o autor.

Na Figura 27 pode ser observada a diferença de desempenho entre as duas técnicas de classificação com relação à taxa de acurácia.

Figura 27: Taxa de acurácia para as técnicas de classificação RNA e K-MEANS.



Fonte: o autor.

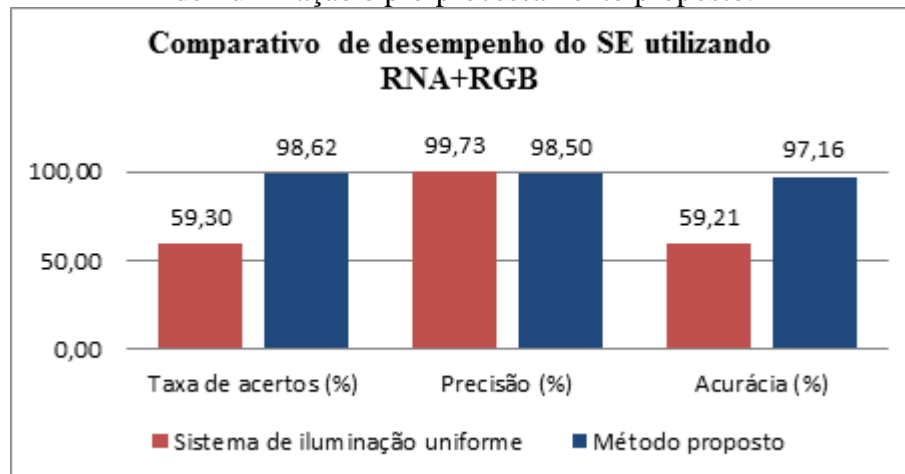
Os erros obtidos durante os experimentos podem ser classificados em dois grupos. O primeiro grupo refere-se aos erros cometidos pelo SE no processo de detecção dos grãos de feijão. Neste caso, alguns grãos deixaram de ser detectados e outros foram detectados mais de uma vez por possuírem tamanhos muito distintos do restante da amostra.

O segundo grupo refere-se ao processo de classificação dos grãos de feijão. Neste caso, os erros cometidos pelo SE estão relacionados com o problema de sincronismo da iluminação e a aquisição das imagens, e também por haver grãos de feijão carioca com tonalidades muito escuras ou com grandes estrias, que foram classificados erroneamente como mulatos.

Nos experimentos realizados, o tempo gasto pelo SE para o processamento de uma amostra variou entre 10 e 15 segundos.

Por fim, é demonstrado na Figura 28, um comparativo de desempenho obtido pelo processo de classificação realizado a partir de RNA e sistema de cores RGB, considerando o esquema que utiliza somente a iluminação uniforme, que é o método comumente empregado pelos SRPV disponíveis na literatura, e o método proposto neste trabalho, que envolve o esquema de iluminação com as lâmpadas dicróicas para geração das sombras em diferentes direções durante a aquisição das imagens, além das técnicas de pré-processamento, que permitiram detectar os grãos de feijão mesmo quando estavam muito próximos ou encostados uns aos outros.

Figura 28: Comparativo de desempenho do SE utilizando RNA+RGB com e sem o esquema de iluminação e pré-processamento proposto.



Fonte: o autor.

Destaca-se que a taxa de precisão não leva em conta os falsos positivos, ou seja, não contabiliza as classificações realizadas incorretamente. Dessa forma, mesmo que esta taxa tenha sido superior para a classificação realizada considerando o sistema de iluminação uniforme, é possível afirmar que houve um ganho na qualidade da classificação quando utilizado o método proposto neste trabalho, visto que a taxa de precisão manteve-se alta, sem que as taxas de acertos e acurácia tivessem sido impactadas com baixos resultados.

5. CONCLUSÕES

Neste trabalho foi desenvolvido um SE voltado para a inspeção da qualidade visual de grãos de feijão, capaz de classificá-los de acordo com a coloração de suas películas de maneira robusta, visando proporcionar maior agilidade e confiabilidade neste processo.

Também foi proposto um método para detecção de grãos que elimina a necessidade de distanciamento dos objetos (sementes/grãos) nas amostras a serem analisadas, sendo uma das principais contribuições deste trabalho. Isso porque esta necessidade de distanciamento se caracteriza como uma limitação existente nos trabalhos encontrados na literatura para análise e classificação de sementes e grãos, que dificulta a aplicabilidade das soluções propostas em processos industriais.

Os resultados obtidos para os quatro grupos de experimentos realizados foram considerados satisfatórios, com taxas médias acima de 90%. Os melhores resultados foram obtidos com a classificação por RNA e sistema de cores RGB, com um valor médio acima de 95% para as três taxas avaliadas.

Como resultado final, conclui-se que o SE desenvolvido usando classificação por RNA e sistema de cores RGB tem potencial para ser incorporado ao processo industrial de inspeção da qualidade visual de grãos de feijão, uma vez que é capaz de detectar os grãos mesmo quando estão próximos ou encostados uns aos outros, além de gerar altas taxas de acertos, precisão e acurácia no processo de classificação.

O tempo gasto pelo SE para o processamento de uma amostra ainda não é ideal para que ele possa operar em um processo contínuo de inspeção. Contudo, este tempo pode ser diminuído, por exemplo, por meio da paralelização de algumas tarefas realizadas pelo SE. Para tanto faz-se necessária uma nova frente de investigação acerca do assunto.

5.1 Sugestões para continuidade do trabalho

Para continuidade deste trabalho, pretende-se realizar experimentos com outros tipos de grãos, uma vez que parte dos processos do SE é parametrizável e outra parte é condicionada a treinamento das técnicas de classificação.

Também pretende-se empregar técnicas de computação paralela para reduzir o tempo de processamento do SE.

Por fim, pretende-se realizar experimentos com o SE em um ambiente de produção de uma indústria de alimentos.

REFERÊNCIAS BIBLIOGRÁFICAS

AGGARWAL, A. K.; RATAN, M. *Aspect Ratio Analysis Using Image Processing for Rice Grain Quality*. *International Journal of Food Engineering*, v. 6, n. 5, article 8, 2010.

ALSABTI, K.; RANKA, S.; SINGH, V. *An efficient k-means clustering algorithm*. *Electrical Engineering and Computer Science*, Paper 43, 1997.

ARAÚJO, S. A. **Casamento de padrões em imagens digitais livre de segmentação e invariante sob transformações de similaridade**. 2009. 142 p. Tese (Doutorado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos. São Paulo. 2009.

ARAÚJO, S. A.; SANTOS, D.; BONALDO, M. A. *Redes Neurais Artificiais Aplicadas em Análise de Perfis Profissionais*. IV Congresso Brasileiro de Computação – **CBCComp** 2004.

BAPTISTA, E. A. **Desenvolvimento de um sistema especialista para a otimização do processo de usinagem**. 2000. 196 p. Tese (Mestrado) – Universidade Metodista de Piracicaba. Santa Bárbara d'Oeste. 2000.

BEAL, A. **Gestão Estratégica da Informação: Como Transformar e Informação e a Tecnologia da Informação em Fatores de Crescimento e de Alto Desempenho nas Organizações**. São Paulo: Atlas, 2004. 137 pp.

BROSNAN, T.; SUN, DW. *Improving quality inspection of food products by computer vision: a review*. *Journal of Food Engineering*, v. 61, n. 1, p. 3-16, 2004.

CHENG F.; YING Y. *Machine vision inspection of rice seed based on Hough transform*. *Journal of Zhejiang University. SCIENCE A*, v.5, n. 6, p. 663-667, 2004.

CONCI, A.; AZEVEDO, E.; LETA, F. R. **Computação Gráfica. Vol. 2 - Processamento de Imagens Digitais**. Rio de Janeiro: Elsevier, 2008.

FAWCETT, T.; *An introduction to ROC analysis*. *Pattern Recognition Letters*. V. 27, n. 8, p. 861-874, 2006.

FERNANDEZ, L.; CASTILLERO, C.; AGUILERA, J. M. *An application of image analysis to dehydration of apple discs*. *Journal of Food Engineering*, v. 67, n.1-2, p. 185–193, 2005.

GUNASEKARAN, S. *Computer vision technology for food quality assurance*. *Trends in Food Science & Technology*, v. 71, n. 8, p. 245-256, 1996.

GIL, A.C. **Métodos e técnicas de pesquisa social**. 4 ed. São Paulo: Atlas, 1994. 207 p.

GONZALEZ, J. M. V.; LINUESA, F. A.; GARCIA, F. L. *A ceramic tile inspection system for detecting corner defects*. *Proceedings of the pattern recognition and image analysis*, Castellón, Spain. 2001.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. Massachusetts: Addison-Wesley, 2002.

GUZMAN, J. D.; PERALTA, E. K. *Classification of Philippine Rice Grains Using Machine Vision and Artificial Neural Networks*. **World Conference on Agricultural Information and IT**. Tokio, 2008.

HAYKIN, SIMON. **Redes Neurais: Princípios e Prática**. trad. Paulo Martins Engel, 2 ed. Porto Alegre: Bookman, 2001.

INTEL. **OpenCV - Open Source Computer Vision Library. 2000**. Disponível em: <<http://www.intel.com/technology/computing/opencv/>>. Acesso em: 15 fev. 2012.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. *Artificial Neural Networks: A Tutorial*. **IEEE Computer**, p. 31-44, Mar. 1996.

JAIN, A. K.; DUIN, R. P. W.; MAIO, J. *Statistical Pattern Recognition: A review*. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 1, p. 4-37, 2000.

KANUNGO T. et al. *An Efficient k-Means Clustering Algorithm: Analysis and Implementation*. **IEEE Transactions on pattern analysis and machine intelligence**, v. 24, n. 7, p. 881-892, 2002.

KASABOV, N. K. **Fundations of neural networks, fuzzy systems and knowledge engineering**. 2. ed. Massachusetts: The MIT Press, 1998.

KILIÇ, K. et al. *A Classification system for beans using computer vision system and artificial neural networks*. **Journal of Food Engineering**, v. 78, n. 3, p. 897-904, 2007.

KIM, H. Y. **ProEikon - Rotinas e programas em C++ para processamento de imagens e visão computacional**. São Paulo, 2006. Disponível em: <<http://www.lps.usp.br/~hae/software>>. Acesso em fev. 2012.

LAUDON, K.; LAUDON, J. **Sistemas de Informação Gerenciais**. 9ª ed. São Paulo: Pearson, 429 pp. 2010.

LAURENT et. al. *Digital camera images processing of hard-to-cook Beans*. **Journal of Engineering and Technology Research**. v. 2, n. 9, p. 177-188, 2010.

LEEMANS, V; DESTAIN, M-F. *A real-time grading method of apples based on features extracted from defects*. **Journal of Food Engineering**, v. 61, n. 1, p. 83–89. 2004.

LIU H.;YU X. *Application Research of k-means Clustering Algorithm in Image Retrieval System*. **Proceedings of the Second Symposium International Computer Science and Computational Technology**, p. 274-277, China, 2009.

LIU, H. et al. *Optimizing Machine Vision Based Applications in Agricultural Products by Artificial Neural Network*. **International Journal of Food Engineering**. v. 7, n. 3, Pages –, ISSN (Online) 1556-3758, 2011.

LOCHT, P.; THOMSEN, K.; MIKKELSEN, P. *Full color image analysis as a tool for quality control and process development in the food industry. In 1997 ASAE Annual International Meeting*, Paper No. 973006. St. Joseph, Michigan, USA: ASAE. 1997

MALINOWSKI, E. L. **Um aplicativo para a execução de sistemas especialistas no planejamento e controle da manutenção.** 2012. 77 p. Dissertação (Mestrado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Tecnológica Federal do Paraná. Ponta Grossa. 2012.

MAPA - MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. **Legislação, LEI nº 6.305, e no Decreto nº 93.563, de 11/11/86, com a Instrução Normativa nº 12 vigente, de 28 de março de 2008.** Disponível em: <http://sistemasweb.agricultura.gov.br/sislegis/action/detalhaAto.do?method=consultarLegislacaoFederal>. Acessado em 20 de março de 2012.

MARUTA, R. H. **Um novo algoritmo de granulometria com aplicação em caracterização de nanoestruturas de silício.** 2011. 80 p. Dissertação (Mestrado em Engenharia Elétrica) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos. São Paulo. 2011.

NOVINI, A. *The latest in vision technology in today's food and beverage container manufacturing industry. In Food Processing Automation IV Proceedings of the 1995 Conference.* St. Joseph, Michigan, USA: ASAE. 1995.

O'BRIEN, J. A. **Sistemas de Informação e as Decisões Gerenciais na Era da Internet.** 2ª ed. São Paulo: Saraiva, 431 pp. 2010.

OSÓRIO, F. S. **Um estudo sobre o reconhecimento visual de caracteres através de redes neurais.** 1991. 302 p. Dissertação (Mestrado). Universidade Federal do Rio Grande do Sul. Curso de Pós-Graduação em Ciências da Computação. Porto Alegre. RS. 1991.

OTSU, N. **A threshold selection method from gray-level histograms.** IEEE Transactions on Systems, Man, and Cybernetics, v. 9, n. 1, p. 62–66, 1979.

PATEL. K. K.; KAR. A.; JHA. S. N.; KHAN. M. A. *Machine vision system: a tool for quality inspection of food and agricultural products. Journal of food science and technology*, v. 49, n. 2, p.123-41, 2012.

PATIL, N. K.; MALEMATH, V. S.; YADAHALLI, R. M. *Color and Texture Based Identification and Classification of food Grains using different Color Models and Haralick features. International Journal on Computer Science and Engineering (IJCSE)*, v. 3, n. 12, p. 3669-3680, 2011.

PAULUS, I; SCHREVEENS, E. *Shape characterisation of new apple cultivars by Fourier expansion of digital images. Journal of Agricultural Engineering Research*, v. 72, n. 2, p. 113–118, 1999.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações.** São Paulo: Thomson, 508 pp. 2008.

PEREIRA, A. C.; REIS, M. S.; SARAIVA, P. M. *Quality Control of Food Products using Image Analysis and Multivariate Statistical Tools. Industrial & Engineering Chemistry Research*. v. 48, n. 2, p. 988-998, 2009.

PIRES, A. C. **Protótipo para inspeção visual automática da qualidade feijão: Concepção, desenvolvimento e análise de custos**. 2012. 88 p. Dissertação (Mestrado) Programa Pós-Graduação em engenharia de Produção, Universidade Nove de Julho, São Paulo, 2012.

RIYADI S. et al. *Shape Characteristics Analysis for Papaya Size Classification. The 5th Student Conference on Research and Developmen*, p. 1-5, 2007.

RODRÍGUEZ-PULIDO et al. *Analysis of food appearance properties by computer vision applying ellipsoids to colour data. Computers and Electronics in Agriculture*. v. 99, p. 108-115, 2013.

ROSARIO, C. R.; Frozza; KIPPER, L. M.; SCHREIBER, J. N. C. Estudo comparativo entre sistema especialista (se) probabilístico e não probabilístico na gestão de conhecimento: estudo de caso em uma empresa metalúrgica do ramo de embalagens metálicas. In: **Enegep**, 2011, Belo Horizonte. XXXI Encontro Nacional de Engenharia de Produção, 2011.

SAVAKAR D. *Identification and Classification of Bulk Fruits Images using Artificial Neural Networks. International Journal of Engineering and Innovative Technology*. v. 1, n. 3, 2012.

SAVAKAR D. G., ANAMI B. S. Recognition and Classification of Food Grains, Fruits and Flowers Using Machine Vision. *International Journal of Food Engineering*. v. 5, n.4, article 14, 2009.

SUN D-W. *Inspecting pizza topping percentage and distribution by a computer vision method. Journal of food Engineering*, v. 44, n. 4 p. 245–249, 2000.

TATIBANA, C. Y.; KAETSU, D. Y. Redes Neurais. **GSI – Grupo de Sistemas Inteligentes**. Disponível em: <<http://www.din.uem.br/ia/neurais/>>. Acessado em: 10/07/2013.

TIMMERMANS, A. J. M. *Computer vision system for online sorting of pot plants based on learning techniques. Acta Horticulturae*, 421, 91–98. 1998.

TOU, J. T., GONZALEZ, R. C. *Pattern Recognition Principles*. Addison-Wesley Publishing Company, Massachusetts, 1981.

VENORA G.; GRILLO O.; RAVALLI C.; CREMONINI R. *Identification of Italian landraces of bean (Phaseolus vulgaris L.) using an image analysis system. Scientia Horticulturae*, v. 121, n. 4, p. 410–418, 2009.

YING, Y.; JING, H.; TAO Y.; ZHANG N. *Detecting stem and shape of pears using fourier transformation and an artificial neural network. Transactions of the ASAE*, v. 46, n. 1, p. 157-162, 2003.

WANG, H-H; SUN, D.-W. *Correlation between cheese meltability determined with a computer vision method and with Arnott and Schreiber. **Journal of Food Science**, v. 67, n. 2, p. 745–749. 2002.*

ANEXO A: TRABALHO PUBLICADO EM ANAIS DE CONGRESSO.

PIRES, A. C.; SANTANA, J. C. C.; PESSOTA, J. H.; ARAÚJO, S. A. *Implementation of a prototype for automatic beans selection* In: ***XIX International Conference on Industrial Engineering and Operations Management (ICIEOM 2013)***, Valladolid. Proceedings of the ICIEOM 2013, 2013. v. 1. p. 1-8.

ANEXO B: TRABALHO PUBLICADO EM PERIÓDICO.

ALVES, W. A. L.; ARAÚJO, S. A.; PESSOTA, J. H.; SANTOS, R. A. B. O. *A Methodology for Sensory Evaluation of Food Products Using Self-Organizing Maps and K-Means Algorithm. Applied Mechanics and Materials*, v. 263-266, p. 2191-2194, 2012

IMPLEMENTATION OF A PROTOTYPE FOR AUTOMATIC BEANS SELECTION

Alexandre Cezar Pires (UNINOVE)

kennedy_xandy@yahoo.com.br

Jorge Henrique Pessota (UNINOVE)

jorgepessota@uninove.edu.br

Sidnei Alves de Araujo (UNINOVE)

saraujo@uninove.br

Jose Carlos Curvelo Santana (UNINOVE)

jccurvelo@uninove.br

In this paper, we discussed about the procedures used for beans classification, the difficulties to maintain the standards required by Brazilian Ministry of Agriculture and developed a low cost prototype, including a computer vision system for automatic selection of beans based on their colors, which according to conducted experiments, could be used in the process of quality inspection of beans.

Keywords: beans, quality, Brazilian standards, computer vision

1 Introduction

Bean is an important source of energy and protein and is present in the daily diet of the Brazilian people. As most of food products, visual properties of beans are important mainly to help the choice of consumers. In Brazil, the quality inspection of food products is usually done manually by following rules and procedures of the Ministry of Agriculture of Brazil (MAB).

Basically, the classification of the beans is carried out manually by extracting a sample of at least 250g of a batch of beans, separating the foreign matter and impurities using a circular sieve with holes of 5 mm of diameter and performing a visual inspection of the sample (after extraction of foreign matter and impurities), to determine the group, class and type based on the operating procedures established by MAB (MAB, 2011).

The main problems encountered in classifying the beans are broken, moldy, burned, crushed, damaged by insects (chopped), sprouted, wrinkled, stained, discolored or damaged by various causes (MAB, 2011).

Other defects include grains and seeds of other plant species found in the analyzed sample, plant debris and foreign bodies. It is considered as foreign matter all the particles become retained in the sieve screens and those of the sample withdrawn normally parsed grain (MAB, 2011).

In addition, the defects of beans can be classified in two groups: light or severe, depending on the severity that makes unfeasible the sale of the product. The beans of the groups I and II (most consumed by Brazilian people) may have three ranks numbered from 1 to 3, according to the quality and maximum tolerances of defects, such as moisture content, impurities and foreign matter found in beans that cannot pass the maximum allowed (MAB, 2011).

Since 1975 until today the Brazilian laws has been improved the classification of the beans. Making a comparison only between a Decree of 1986 and Normative Instruction of 2008, it is possible to verify higher stiffness with respect to the beans classification. Among them, we can cite:

- In all classes and the two groups of the beans, the amount of grains that needs to be the same botanical species in a sample became closer to 100%, making the control almost impossible to be done manually. In group I, this percentage was increased from 95 to 97% and in group II from 80 to 90%.
- The percentage of humidity in the grains was decreased from 15 to 14%;
- The percentage of impurities and foreign matter was decreased from 4% to 2%;
- In 1986 were considered five types in the classification of beans. Today only three are considered.
- The found defects in beans become classified as severe or light and depending on how it is classified the bean cannot be commercialized;
- If the percentage of dead insects exceeds 0.60% beans are disqualified.

According to the points discussed above it is clear that the degree of quality rating has become more difficult to achieve, especially when the process is done manually. This opens the way for studies on alternatives that improve the visual inspection process of Brazilian beans.

According Kiliç (2007) the manual inspection processes applied for food products, including beans, present the inconvenience of time consuming, high cost, parallax errors and the difficulty of standardization of results. Nevertheless, the human sensory detection capabilities can be easily affected by environmental and personal factors (PESANTE-SANTANA and WOLDSTAD, 2000).

In this scenario, it is desirable automated processes that can bring benefits such as cost reduction, standardization of results, prevention of recalls and preservation of the physical integrity of workers. Thus, several researchers have proposed systems for automatic visual inspection tasks (ROSENBERG et al., 2004; KILIÇ et al., 2007; OLIVEIRA et al., 2008; PEREIRA, 2009; VENORA, 2009; MENDOZA et al., 2010; ANAMI and SAVAKAR, 2010; IQBAL, 2011; CASTRO et al., 2012; PEREA et al., 2012).

However, in the literature review, only the papers of (KILIÇ, 2007; VENORA, 2009; LAURENT, 2010) addressing the development of automated systems for classification of beans were found. Nevertheless, these works consider only beans that are not commonly consumed in Brazil.

It is noteworthy that automatic processes to deal with Brazilian beans should consider the operating procedures standardized by MAB. Thus, besides a brief presentation of the procedures for classification of beans and the difficulties to maintain the standards required by MAB, we present the proposed low cost prototype which includes a computer vision system (CVS) applied to visual quality inspection process of beans.

2 Materials and methods

2.1 Proposed low cost prototype

To build the conveyor belt, used in the experiments, it were used two pressure rollers recovered of without using printers (Figure 1a), which contain bearings fixed in their terminals. They were crossed by two long bolts that serve as axes (Figure 1b). In each axis a machining operation was required before fixing them at the bearing holders.

For building of structural part of conveyor belt (chassis), discarded wood was used as raw material. The dimensions adopted for conveyor belt were 1,500 mm of length and 250 mm of width. After the construction and assembly of the chassis, the rollers were fixed to bearing holders attached to the chassis (Figures 1c and 1d). In the last step of building of automated conveyor belt it were built the electronic control circuit of the belt, which is connected to PC parallel port and is responsible for controlling the motor drive belt and the lighting.

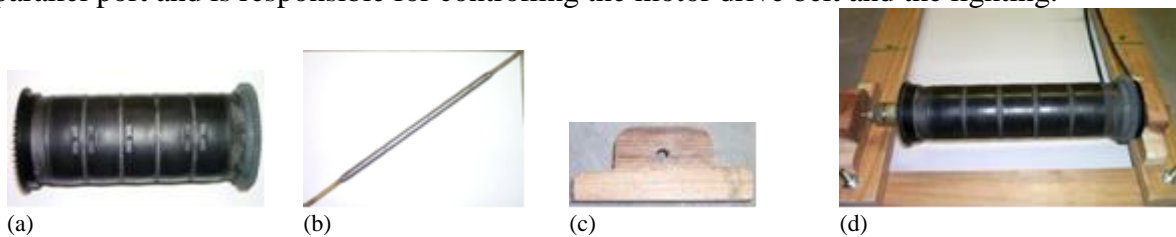


Figure 1: Materials used for building the conveyor belt

2.2 Proposed CVS algorithm

The CVS proposed in this paper was implemented in C/C++ using libraries for Image Processing and Computer Vision ProEikon (KIM, 2006) and OpenCV (INTEL, 2000). We conducted some experiments in which the CVS should be able to recognize most consumed beans in Brazil, i.e. mulatto, black and carioca beans (Figure 4), according to their color.

The input of CVS is an RGB color image, as illustrated in Figure 2a. The first step is to convert this color image into grayscale one (Figure 2b). In the sequence, multilevel Otsu thresholding (LIAO et al., 2001) is applied in grayscale image for grouping the pixels into four classes (C1, C2, C3, and C4), as shown in Figure 2c.

In the following steps the opening operation from Mathematical morphology (GONZALEZ and WOODS, 2002), is applied in the image represented by pixels of C1 and C2 (Figure 2d) and then small

connected components (<50 pixels) are removed since they represent shadow or noise. The results of these operations are shown in figures 2e and 2f, respectively.

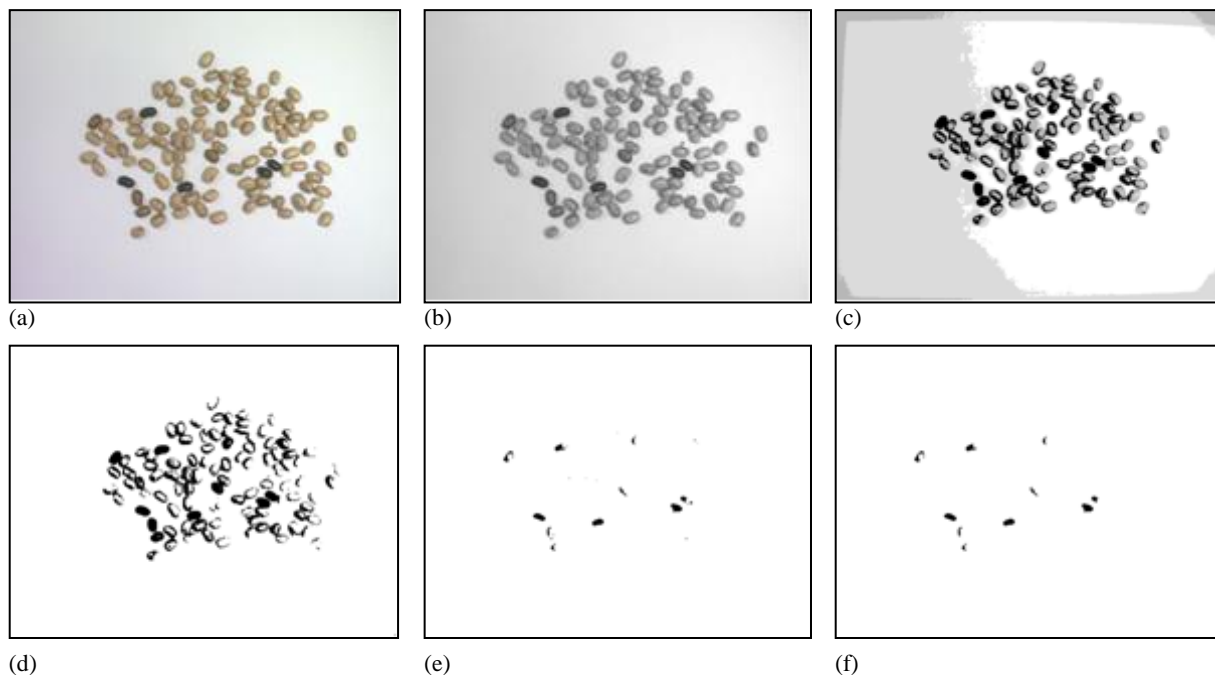


Figure 2: Main steps of the algorithm.

The last steps of CVS consist in calculating the centers of remaining connected components (GONZALEZ and WOODS, 2002) and the use of centers coordinates for counting and labeling the detected grains with a green "X", as showed in Figure 3.

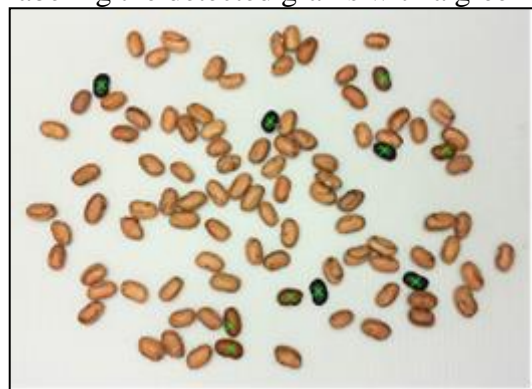


Figure 3: Output image showing detected grains by proposed CVS.

2.3 Experimental setup

In this work we considered most consumed beans for Brazilian people, i.e. mulatto, black and carioca beans (Figure 4), which are purchased in supermarkets in São Paulo, Brazil.

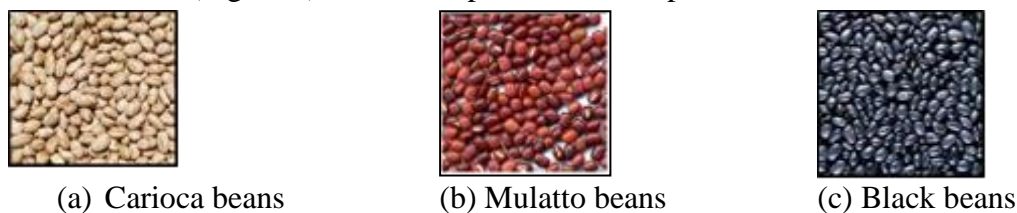


Figure 4: Examples of class/subclass of most consumed Brazilian beans.

In the experiments described in this paper we considered 100 samples of beans, each one containing 100 grains. These samples were divided into 10 subsets, according to the percentage of grains black, mulatto and carioca. The difference among the 10 samples in each subset was the gains positions.

The image acquisition was performed with the use of a Microsoft webcam, coupled to prototype, under the same lighting conditions (fluorescent lamps) and focal length (approximately 40 cm samples of bean), observed in (ANAMI and SAVAKAR , 2010).

In the experiments, the objective of developed CVS was to detect beans not belonging to the subclass carioca in each sample. In summary, for each analyzed image, the algorithm should indicate the number of grains black and mulatto in order to facilitate the calculation of the percentage of the mixture and, therefore, to determine whether the total analyzed portion was approved or rejected.

For each sample of beans the following steps are performed:

- A sample of beans is placed in the conveyor belt;
- The conveyor belt transports the sample of beans to image acquisition box;
- The CVS performs automatically the acquisition and processes the image of the sample;
- Based on image processing, the CVS computes the number of seeds of each class/subclass.

The scheme of working of proposed prototype is illustrated in Figure 5.

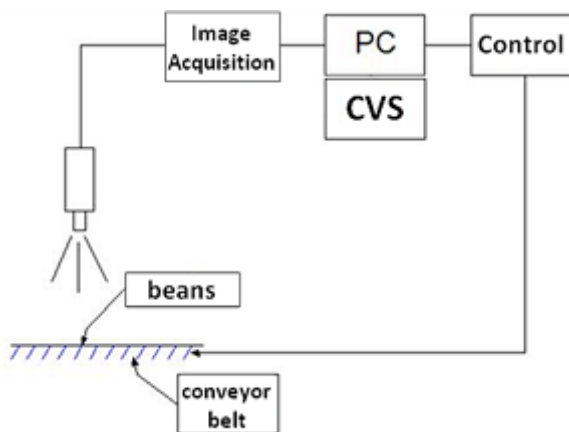


Figure 5: Scheme of working of proposed prototype.

For evaluating the CVS performance hit rate and precision measures (Equations 1 and 2) were used. These measures and others like accuracy, recall-precision curves and ROC curves (FAWCETT, 2005) are widely used to evaluate the performance of CVS algorithms.

$$\text{Recall (hit rate)} = \frac{TP}{TP + FN} \times 100 \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100 \quad (2)$$

where TP is the amount of true positives, FP is the amount of false positives (FP) and FN is the amount of false negative cases.

3 Results and discussions

3.1. Prototype

The complete conveyor belt includes a connection between driver of step motor and the used PC to allow the direction and speed control. It also has a wooden box to fix a digital camera to image acquisition (Figure 6). The total cost of the prototype, including material, labor, and fixed costs, was around \$500.



Figure 6: The proposed prototype.

The prototype was very important to evaluate the proposed CVS simulating a real line of production and its performance was considered satisfactory. Currently, we are working to make it feasible to be used in many other inspection tasks.

3.2. Computer Vision System

The results of experiments with the proposed CVS are summarized in Table 1, which allows us to analyze its performance. Each subset is identified in the first column.

The following columns present, respectively, the amount of grains carioca, black and mulatto in the samples, the total amount of grains to be detected considering all 10 images of the subset, the amount of correct detections (TP), the amount of grains detected incorrectly (FP), the amount of grains not detected (FN) and, finally, the hit rate and precision of the CVS.

The number of TP, FP, FN, hit rate and accuracy were computed based on the amount of grains to be detected.

Table 1: Experimental results obtained with the proposed CVS.

Subset	Amount of carioca beans in each sample	Amount of black beans in each sample	Amount of mulatto beans in each sample	Total amount of grains to be detected	TP	FP	FN	Hit rate (%)	Precision (%)
1	100	0	0	0	0	11	0	100.0	90.1
2	95	0	5	50	48	11	2	96.0	81.4
3	95	5	0	50	50	0	4	100.0	100.0
4	90	5	5	100	93	8	7	93.0	92.1
5	85	5	10	150	129	16	17	86.0	89.0
6	85	10	5	150	134	3	16	89.3	97.8
7	80	10	10	200	171	6	29	85.5	96.6
8	85	0	15	150	143	17	9	95.3	89.4
9	85	15	0	150	147	1	3	98.0	99.3
10	70	15	15	300	277	0	23	92.3	100.0
Total	870	65	65	1300	1192	73	110	91.7	94.2

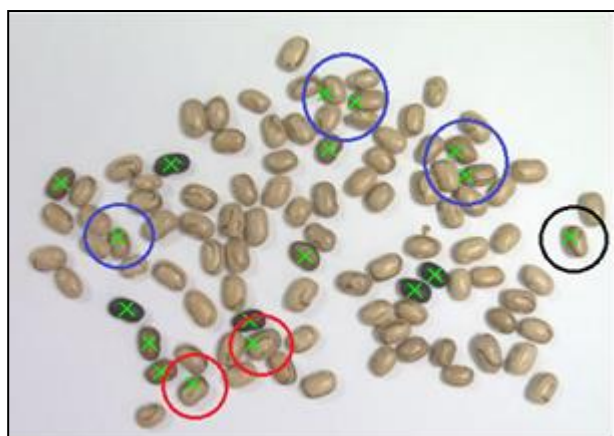
In Table 1 one can see that the overall hit rate of the algorithm was 91.7%. However, the number of false positives decreases its precision to 94.2%.

False positives occurred, in most cases, due to shadows of the grains caused by illumination effects and the existence of longitudinal striations with darker shade in carioca beans, making the algorithm to interpret a few grains as mulatto, as shown in Figure 7a.

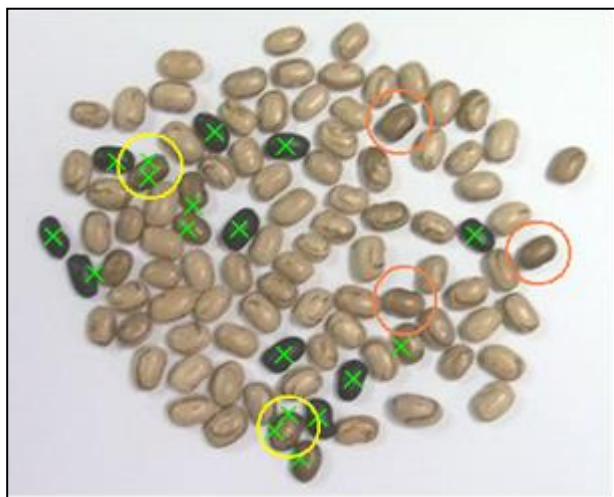
As can be seen in Figure 7b, the direct incidence of light on the sample at the moment of image acquisition causes differences in RGB intensities of the same grain making it to be interpreted by the algorithm as a grain lighter or as two or more grains. This problem leads to increase the number of false positives. In addition, lighting conditions can also cause false negatives errors, as showed in Figure 7b.

It is valid to remember that the value used as threshold for filtering small connected components has a strong influence on the final result of CVS. Very small values can cause false positives decreasing the precision while very large values can cause false negative cases decreasing the hit rate of algorithm. Thus, the suitable choice of this value is important to maximize both hit rate and precision of CVS algorithm.

Another problem is related to the distances among the grains. It was observed that in images where the grains are more separated, the algorithm obtained better results.



(a)



(b)

Figure 7. Output images containing false positive and negative cases. (a) False positives caused by shadow effects (blue circles), longitudinal striations in carioca beans (red circles) and effects of shadow + dark streaks (black circle). (b) False positives (yellow circles) and false negatives (orange circles).

Concerning the time processing, using a microcomputer Pentium4-2.8GHz, the proposed algorithm spent, on average, 1.2 seconds to process an image of 640x480 pixels. This time is suitable for online inspection.

4 Conclusion

In this paper, the procedures used for beans classification and the difficulties to maintain the standards required by the Brazilian Ministry of Agriculture were discussed. We also proposed a low COST prototype for automatic beans selection that could be used to improve the process of quality inspection of beans. It was observed that manual process for beans classification can lead to numerous errors. Brazilian law is rigid in order to ensure a better quality of bean, but became more difficult the classification process that is still done manually. Regarding the experiments using the prototype and CVS proposed, the overall hit rate in detection grains was 91.7%. However, the precision of CVS was decreased to 94.2% because false positive cases. Currently, we are working to make the prototype feasible to be used in practical situations. To this end, we need to improve the CVS become it more robust and able to inspect other types of defects commonly found in Brazilian beans.

Aknowlodgment

The authors would like to thank CNPq (Brazilian National Research Council) for the research scholarship granted to one of the authors.

References

- Anami B S and Savakar D G. (2010). Influence of Light, Distance and Size on Recognition and Classification of Food Grains' Images. International Journal of Food Engineering, 6(2): 1-21.
- Castro J O, Veloso A V, Yanagi Júnior T, Fassani E J, Schiassi L, Campos E T (2012). Estimate of the weight of Japanese quail eggs through *fuzzy* sets theory. Ciência e Agrotecnologia, Lavras, 36(1): 108-116.
- Fawcett T (2006). An introduction to ROC analysis. Pattern Recognition Letters. 27(8): 861-874.
- Gonzalez R C, Woods R E (2002). Digital Image Processing. Massachusetts: Addison-Wesley.
- Intel (2000). OPENCV - Open Source Computer Vision Library. Available in: <<http://www.intel.com/technology/computing/opencv/>>. Accessed at Octubre 05, 2007.

- Iqbal F (2011). Detection of salt affected soil in rice-wheat area using satellite image. *African Journal of Agricultural Research*, 6(21):4973-4982.
- Kiliç K, Boyacı İ H, Köksel H, Küsmenoglu I (2007). A Classification system for beans using computer vision system and artificial neural networks. *Journal of Food Engineering*, 78(1): 897-904.
- Kim, H Y (2006). ProEikon - Routines and Programs in C/C++ for Image Processing and Computer Vision. Available in: <http://www.lps.usp.br/~hae/software>. Accessed at March 05, 2010.
- Laurent B, Ousman B, Dzudie T, Carl M F M, Emmanuel T (2010). Digital camera images processing of hard-to-cook beans. *Journal of Engineering and Technology Research*, 2(9): 177-188.
- Liao P S, Chen T S, Chung P C (2001). A Fast Algorithm for Multilevel Thresholding. *Journal of Information Science and Engineering*, 17(1): 713-727.
- MAB – Ministry of Agriculture of Brazil (2011). Legislation, Law No. 6305 and Decree No. 93563 of 11/11/86, n.12 Instruction Available in: <http://extranet.agricultura.gov.br/sislegis/action/detalhaAto.do?method=consultarLegislacaoFederal>. Accessed at April 10, 2011.
- Mendoza F, Dejmek P, Aguilera J M (2010). Gloss measurements of raw agricultural products using image analysis. *Food Research International*, 43(1): 18-26.
- Oliveira L E S, Cavalin P R, Brito J R A S, Koerich A L (2008). Automatic Inspection of Defects in timber pine using computer vision. *Journal of Theoretical and Applied Informatics*, 15(2): 203-217.
- Perea A J, Meroño J E, Aguilera M J (2011). Hierarchical temporal memory for mapping vineyards using digital aerial photographs. *African Journal of Agricultural Research*, 6(14): 3333-3339.
- Pereira C M G (2009). Machine vision in applications of precision vineyard. Dissertation (Master in Electrical and Computer Engineering), University of Tras-os-Montes and Alto Douro, Portugal.
- Rosemberg C, Emilr B, Laurent H (2004). Calibration and quality control of cherries by artificial vision. *International Journal of Electronic Imaging*, 13(3):539-546.
- Venora G, Grillo O, Ravali C, Cremonini R (2009). Identification of Italian landraces of bean (*Phaseolus vulgaris L.*) using an image analysis system. *Scientia Horticulturae*, 121(4): 410-418.

A methodology for sensory evaluation of food products using Self-Organizing Maps and K-means algorithm

Wonder Alexandre Luz Alves^{1,2,a}, Sidnei Alves de Araújo^{1,b}, Jorge Henrique Pessota^{1,c} and Renato Augusto Barbosa O. dos Santos^{2,d}

¹ Industrial Engineering Post Graduation Program, Nove de Julho University (UNINOVE),
Av. Francisco Matarazzo, 612, 05001-100, São Paulo, SP – Brazil

² Department of Informatics, Nove de Julho University (UNINOVE), Av. Dr. Adolfo Pinto, 109, Barra Funda, São Paulo, São Paulo, SP – Brazil

^awonder@uninove.br, ^bsaraujo@uninove.br, ^cjorgepessota@uninove.edu.br,
^drenatocoy@gmail.com

Keywords: Sensory analysis, Self-Organizing Maps, Kohonen neural network, K-means.

Abstract. Sensory analysis has an important impact on food production since its results can help the understanding of consumers' perceptions about the products. Thus, many methods have been proposed and applied to quantify sensory attributes of food products. In this paper we proposed a methodology, using Kohonen's Self-Organizing Maps and K-means algorithm, to classify food samples through the responses, provided by human evaluators, for their attributes such as aroma, flavor, appearance and texture. Conducted experiments in sensory analysis to determine the acceptance of new gelatins produced from chicken feet and new wines produced from spares of Açaí and Cajá confirm that proposed methodology is suitable for the investigated purpose.

Introduction

Sensory evaluation has been defined as a scientific method used for analyzing food and other materials using the senses of sight, smell, touch, taste, and hearing [1-3]. The sensory evaluation field has grown rapidly in the last decades, mainly because the expansion of the processed-end food and consumer products industries. Nowadays, sensory evaluation becomes an important tool in controlling food products quality.

The sensory analysis allows measuring sensory attributes of the products or even to determine whether the differences in the products are detected and accepted by the consumer or not. In product development or quality control, understanding, determination and evaluation of sensory characteristics of the products becomes important in many situations.

For this reason, many methods and tools have been proposed and applied to quantify the sensory attributes of products, for example the ones proposed by [2-6]. The resulting data are important for demonstrating the sensory evaluation influences in business decisions and to ensure the success of products understanding the perceptions of the consumers about them [1,2].

Tools for decision make in sensory analysis can be improved by using Artificial Intelligence techniques focused in data classification and clusterization such as Kohonen's Neural Network and K-means algorithm.

Kohonen's Neural Network or Self-organizing Map is composed of two layers (input and output) and employs an algorithm for unsupervised learning to translate the similarities of the patterns presented in the input layer in relations of distance between the neurons that compose its output layer [7,8]. It works basically as follows: when a pattern is presented to the network input layer, a neuron of the output layer is chosen to represent this pattern by means of a competitive process. During the training phase, the network increases the similarity of the chosen neuron and their neighbors to the pattern presented in the input layer. Thus, it is constructed a topological map in which the output layer neurons that are topologically close respond similarly to input patterns with similar characteristics.

K-means is a well known clustering algorithm. It consists of a method for clustering analysis whose goal is to partition a set of samples into k clusters in which each sample belongs to the cluster with the nearest mean [9].

In this paper, we proposed a methodology based on Self-Organizing Maps (SOM) and K-means algorithm, for classifying food samples through the responses provided by human evaluators for their attributes such as aroma, flavor, appearance and texture. In order to validate the methodology we also conducted experiments applying it in sensory analysis to determine the acceptance of new gelatins produced from chicken feet and new wines produced from spares of Açai and Cajá.

Proposed Methodology

Let $H \subset Z$ a subset that represents hedonic scales of a factorial design. Thus, the result of sensory analysis can be represented by means of a set $Z = \{(c_i, p_i) \in H^m \times \{1, 2, \dots, P\} : i = 1, 2, \dots, N\}$, such that each element $X_i = (c_i, p_i) \in \zeta$ is a sequence of c_i reviews of m attributes extracted product p_i , made by the evaluator i .

Thus, we analyze local (subset of attributes) and global (all attributes) similarities from the result of sensory analysis, ζ , by means of self-organizing maps. To analyze local similarities, only the results of individual evaluations of each attribute are considered, that is, the sets $Z_j = \{(c_i^j, p_i) \in Z : \text{where } j \text{ is an index of vector } c_i\}$ representing those results. In the global analysis the whole set Z is considered.

Basically, a Kohonen network can be viewed as a mapping of the input data ζ or Z_j in a topological structure of neurons Y , such that when an input X_i is presented to the network, a neuron $y_j \in Y$ is chosen to represent it by means of a competition process, as described before. In this case the neural network architecture is determined according to dimensionality of sets ζ and Z_j .

After the construction of Kohonen's maps (locals and global) their topological structures are analyzed in order to examine the similarities between data sets ζ or Z_j .

The choice of the optimum number of clusters of input data may be determined by a human expert in sensory analysis or by an algorithm. Once determined the number of clusters, we can divide the elements of sets ζ or Z_j in G_k partitions using K-means algorithm as follows: we know that each element $X_i = (c_i, p_i)$ is mapped in neuron $y_j \in Y$ of the neural network. Then, the similarity between the element X_i and a given partition G_t (with $1 \leq t \leq k$) is the distance between the neuron y_j and the centroid of the partition G_t in Y . Thus, we can partition the elements of sets ζ or Z_j minimizing the similarity between the elements of the partitions G_k .

The working of proposed methodology is illustrated in Figure 1.

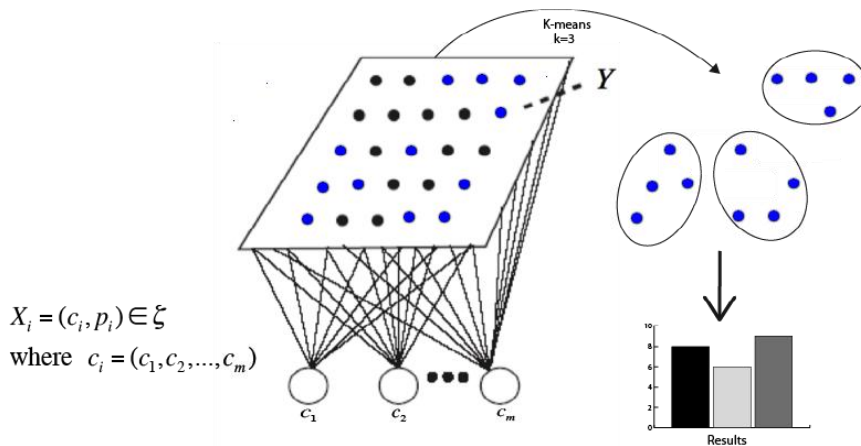


Figure 1. Schematic diagram of proposed methodology.

To show the similarity between the products of sensory analysis, we analyze the probability of occurrence of each product for each G_k partition. Thus, we created a chart so that each bin has only one maximum likelihood product.

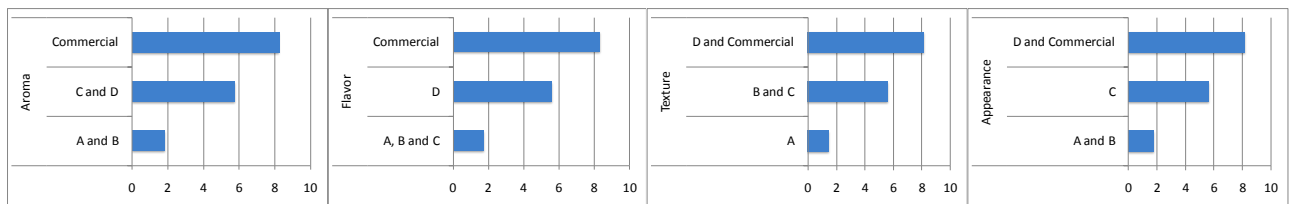
Experimental results

In the following are described two cases in which we applied the proposed methodology:

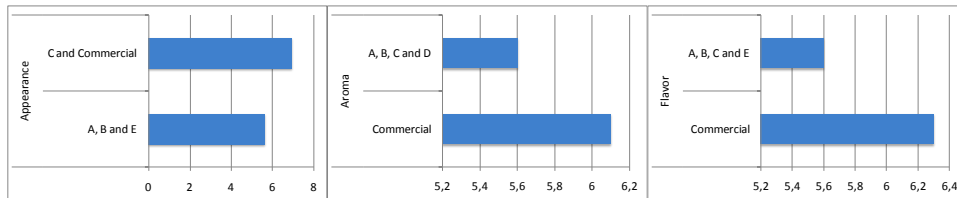
Case 1: It was conducted a sensory analysis to determine the acceptance of new gelatins produced by using chicken feet. In this case it was also considered a gelatin of good acceptance in the commercial market, as reference. Each evaluator i performed a sequence of evaluation c_i (aroma, flavor, texture and appearance) for each gelatin p_i , yielding $X_i = (c_i, p_i) \hat{=} Z$. More details about this experiment can be found in [10].

Case 2: In this case, which is described in detail in [11], it was conducted a sensory analysis to determine the acceptance of new wines produced by using spares of Açai and Cajá. As reference, it was considered a wine grape of good commercial acceptance. Each evaluator i performed a sequence of evaluation c_i (aroma, flavor, and appearance) for each wine p_i composing an element X_i of the result of sensory analysis, that is, $X_i = (c_i, p_i) \hat{=} Z$.

Fig. 2 illustrates the results of sensory analysis obtained in the cases 1 (Fig. 2a) and 2 (Fig. 2b). As can be seen, the number of chosen partitions was three for the first case and two for the second one.



(a) sensory analysis to determine the acceptance of new gelatins



(b) sensory analysis to determine the acceptance of new wines

Figure 2: Obtained results of sensory analysis using the proposed methodology.

It is worth mentioning that the analysis and interpretation of the results depend on an expert in sensory analysis. As mentioned in [10], comparing the proposed methodology with traditional statistical techniques, there is significant gain in the interpretation of the results.

Conclusions

In this paper it was proposed a methodology, using Kohonen's Self-Organizing Map (SOM) and K-means algorithm, for classifying food samples through the responses provided by human evaluators for their attributes. In order to validate the methodology we also conducted experiments applying it in sensory analysis to determine the acceptance of new gelatins and wines produced in Brazil. Based on conducted experiments it was evident that, different from traditional statistical methods, the proposed methodology provided good ratings, even considering widely dispersed data commonly found in the responses of human evaluators in sensory analysis processes. Currently, we are working in the implementation of a software, which employs the proposed methodology, with

interface and resources that facilitate the use of the artificial intelligence techniques described in this work.

Acknowledgements

The authors would like to thank Uninove by financial support.

References

- [1] N.H. Dzung, L. Dzuan and H.D. Tu. The role of sensory evaluation in food quality control, food research and development: a case of coffee study. In: *Proceedings of the 8th Asean food conference*, Hanoi, (2003), p. 862-866.
- [2] M.L. Piana, L. Persano Oddo, A. Bentabol, E. Bruneau, S. Bogdanov and C. Guyot Declerck. Sensory analysis applied to honey: state of the art. *Apidologie* Vol. 35 (2004), p. 26–37.
- [3] M.A. Drake. Sensory Analysis of Dairy Foods. *Journal of Dairy Science* Vol. 90 n. 11 (2007), p. 4925–4937.
- [4] J.C.C. Santana, C.G. Dias, R.R. de Souza and E.B. Tambourgi. Applying of Neural Network on the Wine Sensorial Analysis from Barbados Cherry. *Journal of Food Process Engineering* Vol. 33 (2010), p. 365–378.
- [5] A. Kor and S. Keskin, Quality and Sensory Evaluation for Goat Meat Using Generalized Procrustes Analysis. *Journal of Animal and Veterinary Advances* Vol. 10 (2011), p. 1313–1316.
- [6] E. Ramírez-Rivera, M. Camacho-Escobar, J. García-López, V. Reyes-Borques and M. Rodríguez-Delatorre. Sensory analysis of Creole turkey meat with flash profile method. *Open Journal of Animal Sciences* Vol. 2, (2012) p. 1–10.
- [7] T. Kohonen. The self-organizing map. *Proceedings of the Institute of Electrical and Electronics Engineers*. Vol. 78 (1990), p. 1464–1480.
- [8] S. Haykin. *Neural Networks – A Comprehensive Foundation*. 2nd., Prentice Hall, New Jersey, (1999).
- [9] T. Mitchell. *Machine Learning*. McGraw-Hill, USA, (1997).
- [10] P. F. Almeida; W. A. L. Alves; T. M. B. Farias; J. C. C. Santana. Preparation and sensorial classification of gelatins from chicken feet: correlation using artificial neural network. To appear in: *Información Tecnológica*, 2012.
- [11] A. C. Miranda, W. A. L. Alves, F. H. Pereira, J. C. C. Santana, N. Narain, R. R. Souza. Sensorial classification of euterpe oleracea mart and spondias mombin *L* wines by using kohonen neural network. To appear in: *Proceedings of III International Symposium on Medicinal and Nutraceutical Plants*, 2012.