

Utilização de Técnicas de Processamento de Imagens e Classificação de Padrões no Reconhecimento de Dígitos em Imagens de Medidores de Consumo de Gás Natural*

Use of Image Processing Techniques and Pattern Classification in the Recognition of Digits in Images of Natural Gas Consumption Meters

Julio Cesar Gonçalves¹
Tania Mezzadri Centeno²

Resumo

Este trabalho propõe uma abordagem que emprega técnicas de processamento de imagens e classificação de padrões para o reconhecimento de dígitos apresentados no contador de consumo de medidores de gás natural. Tais imagens são obtidas em campo a partir de condições reais de operação, diferentemente da maioria das abordagens encontradas na literatura, que se baseiam em imagens adquiridas em ambientes controlados. Inicialmente, o contador de consumo é segmentado por técnicas de processamento de imagens. A segmentação é realizada com base no espaço de cor HSL da imagem, diferentemente da maioria dos trabalhos apresentados na literatura, que utilizam imagens em tons de cinza. Em seguida, os dígitos são individualmente segmentados e suas características são extraídas de modo a compor uma base de conhecimento. Essa base serve de apoio para realizar a classificação e o reconhecimento dos dígitos. A metodologia mostrou-se promissora nesse cenário, chegando a alcançar 95% de taxa de acerto no reconhecimento dos dígitos. Apresentou menos de 5% de falha no processo de segmentação do contador de consumo, considerando-se uma base com 903 imagens de medidores de gás.

Palavras-chave: Reconhecimento Ótico de Caracteres (OCR). Redes Neurais Artificiais (RNA). Máquina de Aprendizagem Extremo (ELM). Máquinas de Vetores de Suporte (SVM).

*Submetido em 21/09/2016 – Aceito em 17/04/2017

¹Mestrando em Computação Aplicada Programa de Pós-graduação em Computação Aplicada (PPGCA) da UTFPR, Brasil – juliogoncalves@alunos.utfpr.edu.br

²Doutora em Informática Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial (CPGEI) da UTFPR, Brasil – mezzadri@utfpr.edu.br

Abstract

This work proposes an approach that employs image processing techniques and pattern classification for the recognition of digits displayed in the counter of consumption of natural gas meters. These images are obtained in the field from actual operation conditions, unlike most approaches in the literature that are based on images acquired in controlled environments. Initially, the counter of consumption is segmented by image processing techniques. The segmentation is performed based on the HSL color space of the image, unlike most of the works presented in the literature that uses images in grayscale. Then the digits are individually segmented and their features extracted in order to compose a knowledge base. This base serves as support to perform the classification and recognition of digits. The methodology showed to be promising in this scenario, reaching a 95% success rate in the recognition of digits. Presented less than 5% fault in the process of segmentation of the counter of consumption, considering a database with 903 images of gas meters.

Keywords: Optical Characters Recognition (OCR). Artificial Neural Networks (ANN). Extreme Learning Machine (ELM). Support Vector Machine (SVM).

1 INTRODUÇÃO

O crescimento avançado da tecnologia digital, em conjunto com o desenvolvimento de novos algoritmos, tem permitido um número cada vez maior de aplicações que envolvem processamento digital de imagens (PDI). Paralelo a esse crescimento, tem se tornado cada vez mais comum o uso, no cotidiano, de dispositivos móveis de captura de imagens. É cada vez maior o interesse na extração de informações textuais de imagens do cotidiano capturadas via dispositivos (OBINATA; DUTTA, 2007).

A extração de texto em imagens capturadas em seu cenário natural apresenta questões desafiadoras. Vários fatores influenciam a captura da imagem como o ângulo de captura e a iluminação excessiva, bem como características do próprio texto apresentado, como o tamanho e o tipo da fonte (ZHANG et al., 2002). PDI têm vasta aplicação em sistemas que analisam imagens do cotidiano. Extrair informações de imagens contendo um medidor de consumo de gás enquadra-se nesse contexto.

Uma das tarefas do processo de faturamento em empresas distribuidoras de gás natural é a coleta do volume de consumo. Aplicações voltadas para a execução dessa tarefa se tornam relevantes, uma vez que podem contribuir com a melhoria desse processo em distribuidoras que ainda recorrem à utilização de anotação manuscrita para coletar o valor consumido pelos clientes.

Embora existam atualmente no mercado dispositivos medidores considerados “inteligentes”, que possuem recursos para envio do consumo por telemetria ou que utilizam sistemas de cartões pré-pagos, a forma manuscrita de coleta das informações pode perdurar em concessionárias que não conseguem arcar com o custo da substituição ou da adaptação de um parque de milhares de medidores. Por outro lado, persistir nesse modo de operação requer a digitação dos dados coletados para o sistema de faturamento. Esse tipo de abordagem pode gerar, entre outros problemas, atrasos e incorreções no faturamento e, até mesmo, fraudes durante o transcorrer do processo.

A identificação desse problema motivou esta pesquisa, cujo objeto de estudo foi a extração e o reconhecimento de dígitos apresentados em imagens de equipamentos medidores de consumo de gás natural. Seu propósito é o de buscar uma abordagem robusta e de baixo custo computacional, capaz de ser reproduzida em aplicações voltadas para dispositivos móveis, permitindo que a tarefa de leitura e de reconhecimento dos dígitos contidos no contador de consumo de gás seja executada de maneira eficaz. É disponibilizado ao profissional leitorista uma forma alternativa de coleta de dados ou, mesmo, proporcionando aos clientes que possuam o medidor de gás instalado, no interior de sua residência ou estabelecimento, um meio de informar à empresa distribuidora o valor consumido sem a necessidade de agendar a visita de um profissional leitorista para efetuar a coleta. A possibilidade de a imagem ficar pública poderia ainda passar ao cliente a sensação de que está pagando apenas pelo que consumiu. Busca-se, assim, uma melhoria no processo de coleta de dados, bem como o oferecimento de uma alternativa ao custo de substituição ou adaptação do parque de medidores existente nas distribuidoras.

Outro contexto de aplicação seria na utilização em laboratórios montados pelas distribuidoras para a realização da calibração de medidores.

O objetivo deste trabalho é o de contribuir com o desenvolvimento de um algoritmo que efetue a segmentação baseada no espaço de cores da imagem, diferentemente da maioria das abordagens da literatura que se baseiam em imagens em tons de cinza. Além disso, obter o reconhecimento dos dígitos por meio de três algoritmos distintos de reconhecimento de padrões, comparar seus resultados e contribuir para o avanço do estado-da-arte de detecção e no reconhecimento de dígitos. Espera-se contribuir, também, com a construção de um banco de imagens a ser disponibilizado para que outros pesquisadores possam elaborar novas abordagens para a resolução do problema.

Este artigo toma como base o trabalho publicado pelos autores nos anais do V Simpósio Mineiro de Sistemas de Informação - SMSI 2016. As principais alterações apresentadas neste artigo, em relação ao trabalho outrora publicado, tratam da discussão do método utilizado para gerar o mapeamento de características de cada dígito. Ademais, aborda, também, a discussão dos resultados da validação do método utilizado para reconhecer os dígitos, tanto no que diz respeito ao reconhecimento dos dígitos de forma individual, quanto ao reconhecimento da leitura completa apresentada no medidor.

A sequência deste artigo apresenta, na Seção 2, as pesquisas envolvendo o problema do reconhecimento ótico de caracteres (OCR) em instrumentos de medição. A Seção 3 apresenta a metodologia e as técnicas propostas neste trabalho para solucionar o problema identificado. A Seção 4 apresenta os experimentos e os resultados obtidos, seguindo a metodologia proposta na seção anterior. Concluindo, a Seção 5 apresenta as considerações finais e as sugestões de trabalhos futuros.

2 TRABALHOS RELACIONADOS

Apesar de se apresentar como uma atividade de OCR tradicional, a identificação de texto em objetos, como no caso de medidores, pode apresentar peculiaridades decorrentes de ruídos adquiridos na aquisição da imagem. Esses ruídos podem ser resultantes de diversas fontes, como a falta ou o excesso de iluminação do ambiente, tipo e resolução do sensor da câmera, sombras, oclusões, entre outras interferências. A apresentação dessas peculiaridades pode atrapalhar a interpretação e o reconhecimento de informações dispostas na imagem, inviabilizando a utilização de soluções tradicionais de OCR, como pode ser observado na pesquisa de Nodari e Gallo (2011), que utilizou uma biblioteca padrão de OCR denominada *Tesseract* (SMITH, 2007), obtendo uma taxa de 58% de acerto para o reconhecimento dos dígitos. O resultado, segundo o autor, ressaltou a complexidade da tarefa de OCR neste domínio de aplicação.

Trabalhos encontrados na literatura relatam aplicações para o reconhecimento de placas de automóveis (GAZCÓN et al., 2012; WEN et al., 2011), bem como aplicações para reconhecer textos em objetos específicos, como no caso de rótulos apresentados em garrafas de bebidas

(ARRIGHI et al., 2012). Além disso, a literatura apresenta trabalhos que abordam o reconhecimento de dígitos apresentados em medidores de consumo de água, gás e energia elétrica. Com respeito ao processamento das imagens de medidores, os trabalhos apresentam o emprego de variadas técnicas, dentre as quais a transformação da imagem em tons de cinza (EDWARD, 2013), segmentação por limiarização (ALEXANDRIA, 2004; ZHAO et al., 2009), morfologia matemática (LEI et al., 2010) e rotulação de componentes conexos (OLIVEIRA et al., 2009). Para o reconhecimento dos dígitos, técnicas como Redes Neurais Artificiais (RNA) (ZHAO et al., ; VANETTI et al., 2012), classificação baseada na distância mínima (NAVA-ORTIZ et al., 2011), correlação cruzada (BELAN et al., 2013), entre outras, são utilizadas. São tomadas como base diversas características das imagens dos dígitos, tais como histograma de projeções de pixels (CASTELLS-RUFAS; CARRABINA, 2006), tipo de traços retos ou curvos no dígito (SHU et al., 2007), momentos invariantes (BAI et al., 2010), dentre outras.

3 MÉTODOS PARA SEMENTAÇÃO E RECONHECIMENTO DOS DÍGITOS

A metodologia apresentada neste trabalho tem como propósito a obtenção de um algoritmo que processe imagens de medidores, localize o contador do volume de gás consumido, identifique os dígitos significativos e faça o reconhecimento desses em formato texto. O algoritmo foi projetado para trabalhar com imagens de medidores que possuam as seguintes características: apresentem 4, 5, 6 ou 7 dígitos significativos no contador (dígitos de 0 a 9), bem como apresentem dígitos significativos em cor clara sobre fundo escuro e dígitos descartáveis em cor clara sobre fundo vermelho.

As imagens devem ser capturadas do cotidiano, conforme mostra a Figura 1, em situações reais encontradas por profissionais que efetuam regularmente a leitura, porém sem apresentar obstruções ou ruídos que impossibilitem o seu reconhecimento. Nessa mesma figura, é exibido um padrão encontrado na região de interesse (ROI) que representa a área do contador, formada por dígitos brancos aplicados sobre uma área preta conectada a uma área vermelha. Isso motivou a opção pela segmentação da ROI com base na imagem colorida, diferentemente da maioria dos trabalhos correlatos encontrados na literatura que iniciam o trabalho utilizando imagens em tons de cinza.

Figura 1 – Imagens de medidores com destaque para a característica de apresentação da cor vermelha na área na qual são apresentados os dígitos



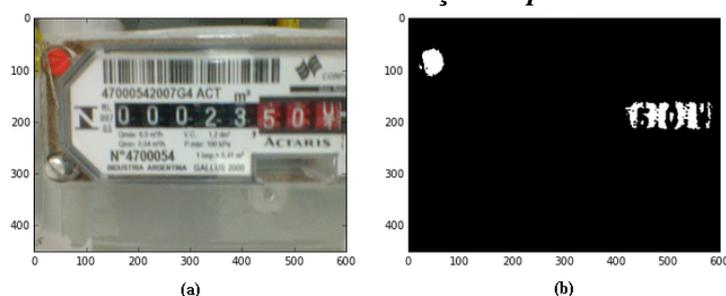
Fonte: Elaborada pelos autores

Nos itens subseqüentes, são descritos os passos que compõem o método proposto para a construção do algoritmo e a obtenção dos resultados desejados.

3.1 Segmentação da região de interesse

A abordagem proposta trata do uso do espaço de cor. Com a conversão da distribuição de valores RGB da imagem para a distribuição de valores HSL, é gerada uma imagem binária (limiarizada), na qual sejam observadas porções que indiquem a presença de *pixels* vermelhos (ou próximos de vermelho). Embora a cor vermelha possa ser obtida por meio da combinação dos valores de RGB, a opção pelo espaço de cores HSL se mostrou mais intuitiva devido à possibilidade de se identificar os *pixels* vermelhos conforme sua luminosidade e saturação. Obtém-se, assim, um tom de vermelho mais forte ou mais fraco de acordo com o ajuste do canal de luminosidade ou mesmo um vermelho mais sujo, apagado ou em sua forma integral, ajustando-se o canal de saturação, como se fosse possível identificar a quantidade de vermelho apresentada em determinado *pixel*. A análise é feita no canal H, que pode possuir tanto um valor consideravelmente baixo quanto um valor consideravelmente alto, bem como nos canais S e L que não poderão conter um valor assim tão baixo. Baseado em experimentos aplicados à base de imagens, foram definidos os seguintes limiares para a filtragem do espaço HSL: $H < h_{inf}$ ou $H > h_{sup}$, $S > s_{min}$ e $L > l_{min}$. A definição destes e de outros parâmetros relacionados à metodologia é tratada mais adiante na seção de experimentos. A Figura 2 mostra o resultado da aplicação da limiarização.

Figura 2 – Imagens identificando ocorrências de *pixels* vermelhos (a) Original (b) Limiarizada com identificação de *pixels* vermelhos



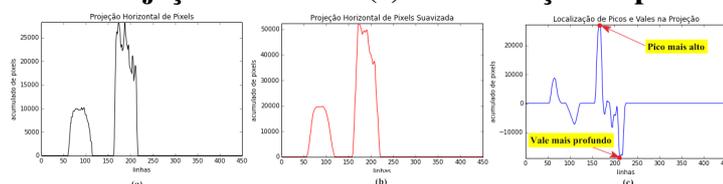
Fonte: Elaborada pelos autores

Baseado na imagem resultante da limiarização gera-se um vetor com a sua projeção horizontal de *pixels*, no qual cada elemento recebe a quantidade de *pixels* vermelhos identificados em cada linha da imagem limiarizada. Para atenuar as variações que ocorrem entre os valores contidos em cada elemento do vetor, aplica-se um filtro de suavização. O filtro consiste em atribuir um novo valor a cada elemento do vetor, com base no cálculo da média dos valores de seus vizinhos, sendo a vizinhança de v_{esq} elementos à esquerda e v_{dir} elementos à direita. O objetivo do filtro é facilitar o processo de identificação de picos e vales presentes na projeção horizontal de *pixels*. Para a localização dos picos e vales, aplica-se um cálculo, levando-se em conta uma vizinhança de três elementos à direita e três à esquerda do valor analisado no vetor de projeções agora suavizado. O novo valor do elemento é obtido por meio da subtração do valor encontrado no seu terceiro vizinho mais à direita pelo valor do seu terceiro vizinho mais à esquerda. Os valores que resultam positivos identificam os picos no vetor de projeções e os negativos identificam os vales. As medidas das vizinhanças, tanto para a suavização quanto

para a obtenção dos picos e vales, foram obtidas com base em experimentos aplicados à base de imagens.

A Figura 3 exibe os gráficos resultantes da projeção horizontal, projeção suavizada e identificação de picos e vales relativos à imagem exibida na Figura 2b.

Figura 3 – Projeção horizontal e identificação de picos e vales (a) Projeção horizontal (b) Projeção suavizada (c) Identificação de picos e vales



Fonte: Elaborada pelos autores

A combinação da identificação do pico mais alto com o vale mais profundo encontrado a sua frente, irá delimitar a altura da ROI (o pico identifica a linha inicial e o vale a linha final). Se a altura encontrada for menor que uma altura mínima, busca-se a combinação com o próximo vale mais profundo até que a altura h_{roi_min} da ROI seja atingida.

Na sequência, segmenta-se a imagem resultante do processo de limiarização com base nos limites de altura encontrados para a ROI. Percorre-se a imagem da direita para a esquerda, analisando-se a projeção vertical (quantidade de *pixels* vermelhos identificados em cada coluna) até encontrar o limite inicial da região vermelha, o que irá delimitar o limite direito da ROI (final da região que contem o contador de consumo). Com os limites superior, inferior e direito definidos, aplica-se uma tolerância de *pixels* acima e abaixo dos limites de altura, procurando evitar a segmentação abrupta ou o corte da ROI em imagens que apresentem captura com leve inclinação. Feito este ajuste, segmenta-se a imagem colorida com base nos três limites obtidos e transforma-se o resultado dessa segmentação em tons de cinza. Toma-se o resultado e realiza-se a limiarização da imagem pelo método de Otsu (1979), executando-se análise semelhante à anterior (projeção vertical) para encontrar o limite esquerdo e segmentar a ROI em sua largura.

O método de Otsu foi escolhido por ser um método eficiente para aplicações cujas características das imagens apresentem uma separação clara de níveis de intensidade entre o objeto a ser segmentado e o fundo, como é o caso dos dígitos presentes nas imagens analisadas neste trabalho, representando o contador de consumo de gás. Imagens similares foram objeto de análise em Lei et al. (2010) e Oliveira et al. (2009), utilizando o mesmo método e obtiveram resultados satisfatórios. A Figura 4 apresenta a segmentação da ROI, utilizando o procedimento descrito.

Figura 4 – Segmentação da ROI (a) Limite direito da ROI (b) Limite esquerdo da ROI (c) ROI segmentada



Fonte: Elaborada pelos autores

A fim de auxiliar a segmentação individual dos dígitos, aplica-se, no resultado, um operador morfológico de erosão por meio de um elemento estruturante retangular de uma linha

e quatro colunas. A aplicação tem o objetivo de eliminar ruídos e separar os dígitos que estejam conectados às linhas que pertencem ao retângulo englobante que contém o contador de consumo.

3.2 Segmentação individual dos dígitos

A segmentação individual dos dígitos parte da imagem resultante do passo anterior e nela aplica-se a análise de componentes conexos (ACC). Essa análise irá buscar, por meio da rotulação de componentes, a transformação da imagem binária até a obtenção de uma imagem simbólica que represente todos os objetos conectados na imagem binária. Um componente conexo é uma região da imagem na qual existe, pelo menos, um caminho entre qualquer par de *pixels* brancos, tal que o caminho é composto apenas por *pixels* brancos. Basicamente, o procedimento para a rotulação de um componente é o de examinar cada *pixel* juntamente com os seus vizinhos e aplicar um *label* (rótulo) a todos os *pixels* que correspondam ao mesmo componente.

Tendo como base a premissa de que parte dos *pixels* que pertencem aos dígitos compõe a linha central da ROI, analisa-se a área de um segmento central composto de duas linhas abaixo e acima a partir da metade da altura da ROI.

Desse modo, são considerados para análise apenas os rótulos resultantes da ACC que possuam *pixels* pertencentes a esse segmento central, descartando-se os demais rótulos. A Figura 5 mostra o resultado do procedimento descrito acima.

Figura 5 – Segmentação dos dígitos (a) Resultado limiarização (b) Resultado da rotulação e linha central de análise (c) Rótulos selecionados (d) Dígitos selecionados



Fonte: Elaborada pelos autores

A partir dos rótulos selecionados, consideram-se como candidatas a dígitos apenas os rótulos que estiverem dentro de um determinado limite px_dig_min mínimo e px_dig_max máximo de *pixels*.

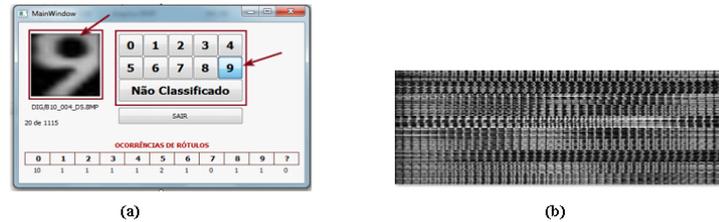
3.3 Reconhecimento dos dígitos

Para a fase de reconhecimento dos dígitos segmentados, foi criada uma base de dados de dígitos rotulados (BDR). Essa base é composta de dígitos padronizados no tamanho de 40x40 *pixels*, que são resultantes do método de segmentação individual de dígitos apresentado neste trabalho. A base é composta de 1220 amostras de dígitos rotulados, sendo 122 amostras de cada dígito (0 a 9), cada qual apresentando 1722 características.

Os rótulos para cada uma das 1220 amostras de dígitos foram atribuídos de forma manual, utilizando uma aplicação desenvolvida para esse fim. A Figura 6a exibe a tela da aplicação

em execução. Por meio disso, um operador humano percorre cada uma das imagens de dígitos e a classifica entre os valores de 0 a 9. De forma automática, a aplicação gera o vetor de 1722 características do respectivo dígito e o vincula ao rótulo atribuído pelo operador.

Figura 6 – Atribuição de rótulos aos dígitos (a) Tela da aplicação (b) Imagem sintética contendo a BDR



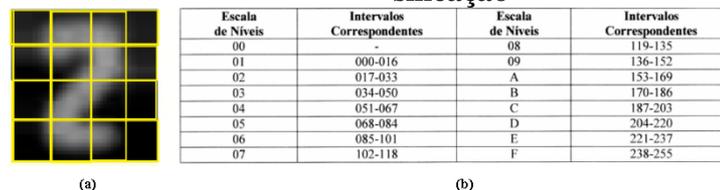
Fonte: Elaborada pelos autores

A conclusão do processo de atribuição de rótulos às 1220 amostras de dígitos gera uma imagem sintética conforme ilustrado na Figura 6b. Nessa imagem, cada linha corresponde a uma amostra de dígito. A primeira coluna de cada linha da imagem é identificada com o rótulo do dígito atribuído pelo operador humano, e as colunas restantes correspondem às características do dígito. Essas características são extraídas pela aplicação das seguintes técnicas: mapeamento do nível de cinza de cada *pixel* (mapa de *pixels*), vetor de comportamento (RUDEK, 1999), *zoning* (PITAS, 2000), projeção horizontal e projeção vertical de *pixels* dos dígitos.

A técnica que utiliza o mapa de *pixels* toma como base a imagem do dígito em tons de cinza e cria um vetor de característica composto por cada um dos valores de nível de cinza apresentado em cada *pixel*. Como cada imagem de dígito foi padronizada no tamanho de 40x40 *pixels*, o resultado da utilização dessa técnica gera um vetor de 1600 características.

A técnica que utiliza o vetor de comportamento toma, também, como base a imagem do dígito em tons de cinza. Realiza a subdivisão da área do dígito em 16 fatias de mesma altura e largura, conforme demonstra a Figura 7a. Calcula a média dos níveis de cinza presentes em cada uma das fatias e classifica o resultado conforme a divisão de intervalos apresentada na tabela exibida na Figura 7b. Por exemplo, caso a primeira fatia do dígito apresente o nível de cinza médio igual a 52, a fatia será classificada no nível 04 da tabela. Desse modo, é gerado um vetor de 16 características para o dígito.

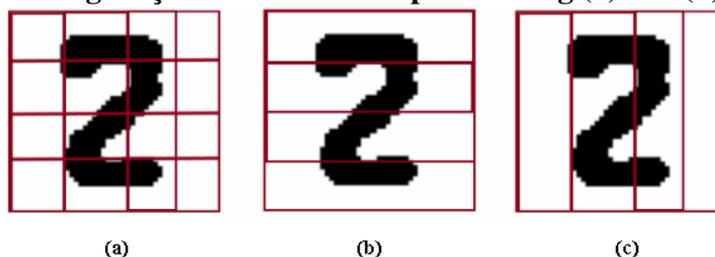
Figura 7 – Geração do vetor de comportamento (a) Dígito fatiado 4x4 (b) Tabela de classificação



Fonte: Adaptado de RUDEK, 1999

De modo similar, a técnica de *zoning* se vale do fatiamento da imagem, porém, toma como base a imagem limiarizada do dígito. São utilizadas três configurações de fatiamento: 4x4, 4x1 e 1x4, conforme ilustra a Figura 8. Calcula-se o número de *pixels* pretos apresentados em cada fatia, gerando um vetor de 24 características para o dígito.

Figura 8 – Configurações de fatiamento para zoning (a) 4x4 (b) 4x1 (c) 1x4



Fonte: Adaptado de PITAS, 2000

As técnicas de projeção horizontal e vertical também tomam como base a versão limiarizada dos dígitos. A primeira gera um vetor de 40 características com base na distribuição de *pixels* pretos apresentados em cada linha da imagem limiarizada do dígito. A segunda gera outro vetor de 40 características baseado na distribuição de *pixels* apresentados em cada coluna. A quantidade de características extraída por cada técnica, bem como o posicionamento dos valores no vetor de características são apresentados na Tabela 1.

Tabela 1 – Mapeamento de características extraídas dos dígitos

Técnica	Característica	Posição inicial no vetor	Posição final no vetor
Mapa de <i>pixels</i>	1600	0001	1600
Vetor de Comportamento	16	1601	1616
Zoning 4x4	16	1617	1632
Zoning linhas	4	1633	1636
Zoning colunas	4	1637	1640
Zoning diagonal	2	1641	1642
Projeção horizontal	40	1643	1682
Projeção vertical	40	1683	1722

Fonte: Elaborada pelos autores

As imagens de medidores utilizadas para obtenção das 1220 amostras de dígitos foram obtidas com base em dois grupos de imagens. O primeiro grupo (G1) foi originado de imagens fornecidas pela Companhia Paranaense de Gás (COMPAGAS). O segundo grupo (G2) foi oriundo de imagens disponibilizadas pelo laboratório de pesquisas italiano Arte-Lab 3³, mesma base utilizada no trabalho de Nodari e Gallo (2011). Vale ressaltar que, no trabalho de Nodari, a segmentação da ROI parte de uma imagem em tons de cinza, sendo que, em alguns casos, o grupo G2 apresenta imagens que são capturadas sem destacar de forma suficiente a parte vermelha do contador de consumo, fato que inviabiliza o método de segmentação da ROI aqui proposto. Sendo assim, foram utilizadas 640 imagens de medidores do grupo G1, originando 930 dígitos individuais e 153 imagens de medidores do grupo G2, originando 290 dígitos individuais. O número relativamente baixo de dígitos, se comparado ao número de imagens de medidores, se deve ao fato de que a segmentação individual gerou quantidades diferentes para cada dígito (0 a 9). Por exemplo, o dígito zero apresentou 1790 ocorrências, enquanto que o dígito nove apresentou 122 ocorrências. Desse modo, para evitar disparidade, foram seleciona-

³Banco de imagens disponível em: <<http://artelab.dicom.uninsubria.it/downloads.html>>, ver item "Gas-meter reading"(acessado em 01/04/2017).

das, aleatoriamente, 122 ocorrências para cada dígito de 0 a 9, totalizando as 1220 amostras que integram a base. Todas as imagens que compõem a base, tanto de medidores quanto a BDR, foram disponibilizadas em repositório *online*⁴, a fim de que outros pesquisadores possam fazer uso em pesquisas futuras envolvendo o tema.

O reconhecimento ficou a cargo de três classificadores: *K-Nearest Neighbours* (KNN), *Support Vector Machine* (SVM) e *Extreme Learning Machine* (ELM).

O método para classificação via KNN consiste, resumidamente, em atribuir uma classe a um elemento desconhecido usando a classe da maioria de seus k vizinhos mais próximos, segundo uma determinada distância (no espaço de atributos). Dada determinada amostra de uma classe desconhecida, calcula-se a sua distância perante todas as amostras da base de conhecimento, selecionando-se as k amostras mais próximas. Dentre essas amostras, escolhe-se a classe que seja maioria entre elas. Se houver empate, utiliza-se a classe mais próxima das maiorias empatadas. Essa técnica apresenta um custo computacional elevado para bases de conhecimento com um grande número de amostras. Isso se deve à necessidade do cálculo da função de distância da amostra desconhecida para todas as amostras contidas na base de conhecimento.

SVM é um método de aprendizagem de máquina desenvolvido por Vapnik (2013), o qual propõe a criação de um hiperplano como superfície de decisão. Esse hiperplano deve apresentar a separação ideal entre duas classes, maximizando a distância entre os pontos mais próximos da classe e o hiperplano. Esses pontos são denominados elementos de borda e os vetores de suporte são as distâncias entre os elementos de borda e o hiperplano. Baseia-se na Teoria de Aprendizagem Estatística como estratégia de treinamento e na Teoria de Otimização para encontrar o hiperplano (SANTOS, 2002). Possui a vantagem de ser capaz de trabalhar com um grande volume de amostras e da rapidez na classificação, embora o tempo de treinamento possa apresentar custo computacional considerável para bases com elevado número de amostras.

O método ELM (HUANG et al., 2007) é derivado da visão atual de RNA como aproximadores universais. Utiliza redes *Single-hidden Layer Feedforward Network* (SLFN), consideradas um caso especial das redes neurais MLP, por apresentarem uma única camada oculta (SILVESTRE, 2015). Destaca-se devido ao baixo custo computacional para treinamento da rede, sendo que sua essência está no uso de pesos sinápticos aleatórios na camada oculta. Desse tipo de configuração, origina-se um modelo linear para os pesos sinápticos da camada de saída da rede. Esses pesos são calculados de forma analítica por meio de uma solução de mínimos quadrados (HUANG et al., 2007). O método não exige parametrizações, que são características de controle de RNA, como o ajuste manual da taxa de aprendizagem e do número de iterações, bem como o ajuste do desempenho de critérios de parada. A aplicação da ELM também resolve o problema dos mínimos locais em relação ao algoritmo back-propagation, já que o método ELM não é baseado nas teorias de gradiente descendente.

Para a avaliação do desempenho dos algoritmos, utilizou-se o método de validação *K fold*. O método consiste em aplicar partições no conjunto de amostras, combinando-as entre

⁴Banco de imagens utilizado neste trabalho, disponível em: <<https://github.com/jcgcwb/gas-meter-ocr>> (acessado em 01/04/2017).

treino e teste (STONE, 1974). As amostras são divididas em k partes de igual tamanho, fazendo com que uma das k partes seja selecionada como conjunto de teste e estabelecendo as $k-1$ partes restantes como conjunto de treinamento. O conjunto de teste é alternado em cada passo do processo que será executado k vezes. A taxa de avaliação é calculada com base na média das taxas de erro e/ou acerto obtido em cada iteração.

4 EXPERIMENTOS E RESULTADOS OBTIDOS

Os algoritmos elaborados para o experimento foram desenvolvidos em linguagem de programação *Python 2.7*, utilizando as bibliotecas *PIL* e *ScikitImage* para o processamento de imagens, *Numpy* para o processamento de matrizes e a biblioteca *ScikitLearn* para os classificadores KNN e SVM, todas disponíveis por meio da instalação do pacote de distribuição Anaconda⁵. Para o classificador ELM, foi utilizada a biblioteca *HP-ELM Toolbox*⁶. Os algoritmos foram testados com o apoio de um *notebook* configurado com o sistema operacional Windows 7, 32bits, 4GB de memória RAM e CPU Intel Core2Duo T6600 de 2.20GHz.

4.1 Parametrização do Experimento

Durante o experimento foram testados alguns parâmetros, tanto para a segmentação de regiões vermelhas, tomando como base o espaço HSL de cores da imagem, quanto para a localização da ROI com base na análise da projeção horizontal de *pixels* suavizada. Também foram definidos parâmetros para a seleção dos rótulos originados na ACC, que levam a segmentação individual dos dígitos. A Tabela 2 apresenta os valores definidos como mais adequados para cada parâmetro, após experimentos que foram aplicados à base de imagens.

Com relação ao estudo de parâmetros utilizados na análise do espaço HSL, observou-se que valores diferentes dos apresentados na Tabela 2 identificavam, na imagem, áreas próximas ao amarelo como sendo áreas vermelhas ou, conforme o ajuste, áreas próximas ao preto como sendo áreas vermelhas. Com relação à suavização da projeção horizontal de *pixels*, à medida que a janela de vizinhança é diminuída em relação ao valor apresentado na Tabela 2, o nível de suavização passa a apresentar-se inadequado devido aos picos e aos vales cada vez mais ruidosos. A atribuição de outra combinação de limites de *pixels* para as áreas de rótulos, que passam pela linha central da ROI, faz com que além dos dígitos individuais, sejam também selecionados ruídos que entremeiam os dígitos.

⁵Pacote de distribuição Anaconda disponível em: <<https://www.continuum.io/downloads>> (acessado em 01/04/2017).

⁶Repositório Github da biblioteca HP-ELM disponível em: <<https://github.com/akusok/hpelm>> (acessado em 01/04/2017).

Tabela 2 – Parâmetros definidos no experimento

Fase	Parâmetro	Valor	Descrição
Segmentação da ROI (análise do espaço HSL)	h_inf	<15	Limite inferior para o canal H do HSL.
	h_sup	>330	Limite superior para o canal H do HSL.
	s_min	>30	Limite mínimo para o canal S do HSL.
	l_min	>10	Limite mínimo para o canal L do HSL.
Segmentação da ROI (suavização da projeção horizontal)	v_esq	=10	Tamanho da vizinhança esquerda para suavização da projeção horizontal de pixels.
	v_dir	=10	Tamanho da vizinhança esquerda para suavização da projeção horizontal de pixels.
Segmentação dos dígitos (componentes conexos)	h_roi_min	=25	Altura mínima para a ROI identificada com a análise de picos e vales da projeção horizontal.
	px_dig_min	=10	Limite mínimo de pixels a ser apresentado no rótulo que passa pela linha central da ROI.
	px_dig_max	=70	Limite máximo de pixels a ser apresentado no rótulo que passa pela linha central da ROI.

Fonte: Elaborada pelos autores

4.2 Experimentos com a localização da ROI

O experimento do método proposto para localização da ROI, com base no espaço de cores da imagem, foi executado nos grupos de imagem G1 e G2. A Tabela 3 apresenta as taxas de acerto relacionadas ao método de localização da ROI.

Tabela 3 – Taxas de acerto na localização da ROI

Grupo	N de imagens	ROI Localizada	ROI Não Localizada	Acerto
G1	640	615	25	96,09%
G2	153	135	18	88,24%

Fonte: Elaborada pelos autores

Em uma avaliação visual qualitativa, observou-se que, em média, aproximadamente 92% das 793 imagens processadas tiveram o resultado da detecção da ROI, apresentando o contador de consumo de forma integral. Aproximadamente 8% das imagens apresentaram a segmentação da ROI de forma inadequada.

A Figura 9 ilustra alguns exemplos de imagens que resultaram em falha na segmentação da ROI. A figura apresenta a imagem do medidor com o respectivo resultado da limiarização para fim de localização de *pixels* vermelhos.



Fonte: Adaptado de PITAS, 2000

Foi possível observar que as falhas, em sua maioria, foram decorrentes de problemas com a iluminação do ambiente no qual a imagem do medidor foi capturada, bem como, em alguns casos, pela falta da porção vermelha pertencente ao contador de consumo.

4.3 Experimentos com o reconhecimento dos dígitos

Para o reconhecimento dos dígitos, foram utilizados os grupos de imagem G1 e G2 de forma individual e em conjunto. A validação cruzada em 10 etapas ($k=10$) foi utilizada para medir o desempenho de cada classificador. Para o classificador ELM, foi escolhida uma arquitetura de 1722 unidades de entrada (representando o vetor de características), 80 neurônios na camada oculta e 10 saídas (representando os dígitos de 0 a 9). A função de ativação utilizada foi a sigmóide. O número de neurônios na camada oculta e a função de ativação foram selecionados, empiricamente, com base na avaliação da taxa de acerto no momento do treinamento. Para o classificador SVM, foi escolhido o *kernel* linear com parâmetro de penalidade $C=1$, também escolhido, empiricamente, com base em análise do acerto no treinamento. Para o classificador KNN, foi escolhido $k=3$ após avaliação da taxa de acerto em experimentos com k variando de 1 a 18 (em intervalos de 2). Para o cálculo da distância, foi utilizada a distância Euclidiana.

A Tabela 4 apresenta as taxas de acerto relacionadas ao reconhecimento dos dígitos em cada grupo de imagem, bem como o tempo médio decorrido em segundos, de acordo com cada etapa da validação cruzada.

Os resultados revelam classificadores com taxas de acerto muito próximas, com o KNN apresentando desempenho ligeiramente inferior em comparação aos métodos ELM e SVM, porém com um alto custo computacional em relação a esses dois últimos. Já a classificação por ELM se destaca com um custo computacional bem menor comparado aos demais classificadores.

Tabela 4 – Taxas de acerto na localização da ROI

Validação Cruzada	Classificador ELM			Classificador SVM			Classificador KNN		
	G1	G2	G1+G2	G1	G2	G1+G2	G1	G2	G1+G2
<i>Fold-1</i>	94%	94%	94%	96%	90%	96%	94%	87%	91%
<i>Fold-2</i>	94%	93%	92%	98%	87%	99%	89%	80%	93%
<i>Fold-3</i>	94%	93%	94%	95%	97%	97%	81%	97%	89%
<i>Fold-4</i>	94%	94%	94%	94%	97%	94%	93%	83%	86%
<i>Fold-5</i>	94%	93%	94%	91%	97%	94%	91%	83%	93%
<i>Fold-6</i>	95%	93%	94%	97%	90%	94%	94%	87%	95%
<i>Fold-7</i>	94%	93%	94%	94%	93%	97%	88%	90%	92%
<i>Fold-8</i>	95%	93%	94%	98%	99%	98%	93%	90%	90%
<i>Fold-9</i>	95%	94%	95%	98%	90%	95%	86%	97%	92%
<i>Fold-10</i>	95%	93%	94%	98%	95%	95%	90%	90%	93%
Acerto Médio	94,4%	93,3%	93,9%	95,9%	93,5%	95,9%	89,9%	88,4%	91,4%
Tempo Médio	0,11 s	0,11 s	0,17 s	1,20 s	0,23 s	1,96 s	4,97 s	0,63 s	8,74 s

Fonte: Elaborada pelos autores

A fim de validar o método proposto, foi utilizado um novo grupo de imagens composto por 110 imagens de medidores. Assim como no grupo G1, estas imagens, também, foram fornecidas pela COMPAGAS. Algumas amostras destas imagens podem ser observadas na Figura 10.

Figura 10 – Amostras de imagens do grupo utilizado para validar o método



Fonte: Elaborada pelos autores

Com relação à fase de segmentação da ROI, para o grupo de imagens de validação, o resultado da segmentação foi de 100%. A fase de segmentação individual dos dígitos produziu 544 dígitos, que foram submetidos aos classificadores ELM e SVM na fase de reconhecimento. Devido ao desempenho inferior do classificador KNN em relação ao tempo de processamento, decidiu-se por utilizar apenas os classificadores ELM e SVM na validação do método.

A validação levou em conta o percentual de acerto no reconhecimento de dígitos, bem como o percentual no qual o reconhecimento resultou em leitura correta, parcial ou incorreta. Entenda-se com leitura correta aquela que ocorre quando o resultado do reconhecimento dos dígitos coincide na forma e na sequência exata em que a leitura se apresenta na imagem do medidor. Para o diagnóstico de leitura parcial, leva-se em consideração um erro de tolerância, que admite a possibilidade do dígito mais à direita da leitura original estar em desacordo com o resultado do reconhecimento. Qualquer outro tipo de resultado retornado leva ao entendimento

de ocorrência de leitura incorreta. A Figura 11 ilustra a forma na qual a classificação da leitura pode ser identificada.

Figura 11 – Amostras de imagens do grupo utilizado para validar o método



Fonte: Elaborada pelos autores

A Tabela 5 apresenta, para cada classificador, os resultados que correspondem ao diagnóstico de precisão nas leituras, bem como do reconhecimento de dígitos. Juntamente com cada resultado percentual aparece, entre parênteses, o número de leituras ocorridas de forma correta, parcial ou incorreta, bem como a quantidade de dígitos reconhecidos corretamente.

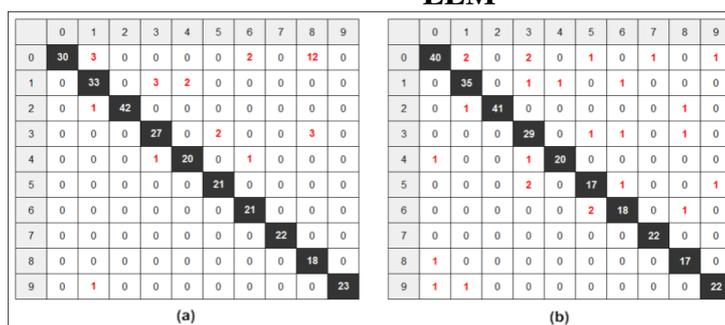
Tabela 5 – Percentuais de diagnóstico de leitura e de reconhecimento de dígitos

Classificador	Leituras	Leituras	Leituras	Dígitos
	Corretas	Parciais	Incorretas	Reconhecidos
SVM	80,91% (89)	1,82 (2)	17,27% (19)	94,30% (513)
ELM	84,55% (93)	4,55 (5)	10,91% (12)	95,04% (517)

Fonte: Elaborada pelos autores

A Figura 12 apresenta a matriz de confusão gerada com base no resultado obtido com a fase de reconhecimento de dígitos utilizando os dois classificadores. Na matriz, as linhas representam os valores previstos para cada dígito, enquanto que as colunas representam o valor retornado por cada classificador na tentativa de reconhecer o dígito. A diagonal principal destaca os acertos encontrados na classificação dos dígitos, enquanto que os demais valores em vermelho representam os erros de classificação observados para cada dígito.

Figura 12 – Matriz de confusão (a) Resultado utilizando SVM (b) Resultado utilizando ELM



Fonte: Elaborada pelos autores

De modo diferente ao ocorrido na etapa de treino e de teste dos algoritmos de classificação, que utilizou uma base de dígitos balanceada e levou em conta apenas o reconhecimento individual dos dígitos, essa etapa de validação levou em conta, ao mesmo tempo, o reconhecimento individual de cada dígito, bem como o reconhecimento da leitura total apresentada em cada imagem de medidor. O fato de ter se levado em conta, ao mesmo tempo, a leitura total do medidor, e de não ser possível ter o controle e o balanceamento da ocorrência de dígitos em cada leitura, impossibilitou que a validação do reconhecimento individual dos dígitos, fosse efetuada em uma base balanceada, conforme ocorrido com o experimento de treino e de teste

dos classificadores. Embora uma base desbalanceada possa produzir efeitos enviesados, o resultado apresentado pelo experimento de validação reforçou o que foi apresentado pelo método *K-fold cross validation* em 10 etapas, com taxas de acerto individual dos dígitos semelhantes entre os classificadores ELM e SVM.

No exercício de validação, também, foi analisado o tempo médio de execução em cada uma das fases do método proposto. Foi investigado o tempo médio decorrido em segundos para segmentar a ROI, segmentar os dígitos e extrair suas características, bem como executar o reconhecimento dos dígitos, tanto utilizando o classificador SVM, quanto o classificador ELM. Foi observado, também, o tempo total decorrido para execução completa do método utilizando cada um dos classificadores. A Tabela 6 apresenta os resultados obtidos.

Tabela 6 – Percentuais de diagnóstico de leitura e de reconhecimento de dígitos

Segmentação ROI	Segmentação Dígitos	OCR via SVM	OCR via ELM	Total com SVM	Total com ELM
0,73 s	0,21 s	2,13 s	0,24 s	3,07 s	1,18 s

Fonte: Elaborada pelos autores

Com as taxas de acerto obtidas no experimento, foi possível confrontar os resultados com alguns casos apresentados na literatura, nos quais os autores destacaram as taxas de acertos relacionadas ao processo de reconhecimento dos dígitos. A Tabela 7 destaca, de acordo com o trabalho de cada autor, as seguintes informações: a taxa de acerto e o tipo de classificação utilizada para o reconhecimento dos dígitos, a presença de controle na captura das imagens (câmera ou outro dispositivo de captura acoplado ao medidor) e o número de imagens de medidores utilizado no experimento.

Tabela 7 – Taxas de reconhecimento obtidas nos trabalhos relacionados

Autor	Controle	Classificação	Acerto	Número de imagens
Castells-Rufas e Carrabina (2006)	Sim	Programação Dinâmica	99%	396
Zhao et al. (2009)	Sim	RNA	95%	500
Bai et al. (2010)	Sim	Momentos de Hu	94%	50
Lei et al. (2010)	Sem relato	Análise de esqueleto	99%	850
Nava-Ortiz et al. (2011)	Sim	Distância mínima	93%	217
Nodari e Gallo (2011)	Não	OCR Tesseract	58%	Sem relato
Zhao et al. ()	Sem relato	RNA	99%	50

Fonte: Elaborada pelos autores

Analisando-se os resultados apresentados na literatura, foi possível concluir que o resultado obtido neste trabalho, de cerca de 95% de taxa de acerto no reconhecimento dos dígitos pode ser considerado promissor. Vale ressaltar que, em boa parte dos casos relacionados na literatura, as imagens de medidores são capturadas em ambientes controlados ou em ambiente de laboratório. Outra comparação pertinente seria com a pesquisa de Nodari e Gallo (2011), a qual obteve resultado de 58% e que, assim como neste trabalho, utiliza imagens capturadas em campo e que apresentam maior conformidade com a situação encontrada pelos profissionais de leitura no seu dia a dia.

5 CONCLUSÃO

Este trabalho teve como objetivo apresentar a metodologia adotada para a construção de um algoritmo envolvendo técnicas de PDI e de reconhecimento de padrões. Esse algoritmo visa obter, em imagens de medidores de gás natural, o reconhecimento dos dígitos apresentados no contador de consumo de gás.

As melhores taxas de acerto foram próximas de 95%, tanto para o classificador SVM quanto para o classificador ELM. No entanto, vale ressaltar que o método que utilizava ELM apresentou um custo computacional consideravelmente menor. Esse é um fator significativo para se determinar a utilização de uma aplicação dessa natureza em dispositivos móveis, num ambiente real, que envolva o processo de faturamento em uma empresa distribuidora de gás natural.

O estudo e sua aplicação se mostraram com potencial promissor, embora o método proposto apresente limitações relacionadas ao reconhecimento de dígitos em transição, ser afetado por variação na escala dos dígitos e ter dificuldade de localizar a ROI em imagens ruidosas devido a pouca luminosidade. Adequações ao processo de segmentação utilizando métodos como o detector de bordas de Sobel, em conjunto com a transformada de Hough, para a correção de imagens que apresentem inclinação, bem como a utilização de outros descritores, como os momentos invariantes de Hu, a fim de lidar com a variação de escala, ou mesmo com a presença de concavidades, furos e traços para tratamento de dígitos em transição, podem ser caminhos para que se alcance melhores taxas de acerto no reconhecimento dos dígitos. Com relação a ruídos apresentados devido à falta de luminosidade, um caminho seria o emprego de um filtro para ajuste de contraste. Além disso, a aplicação de um razoável nível de controle na captura das imagens, como por exemplo, fazer com que o contador de consumo esteja contido dentro de um retângulo aparente presente na tela do dispositivo no momento da captura, pode também ser uma maneira de elevar a taxa de reconhecimento. Seria interessante experimentar o algoritmo com um *hardware* de baixo custo como um celular, possibilitando assim o levantamento da quantidade de leituras em campo obtidas de modo acertado. Essas poderiam ser consideradas como atividades futuras para continuidade desta pesquisa.

REFERÊNCIAS

ALEXANDRIA, Auzuir R. **Sistema de Reconhecimento de Algarismos em Medidores de Energia Convencionais Visando Implementação de Sistema de Fornecimento de Energia Pré-Paga – Seppra**. 2004. Dissertação (Mestrado) — Universidade Federal do Ceará.

ARRIGHI, T. et al. Recognition and classification of numerical labels using digital image processing techniques. In: **XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA)**. Colombia: IEEE, 2012. p. 252 – 260.

BAI, Qiushi et al. Research of automatic recognition of digital meter reading based on intelligent image processing. In: **Computer Engineering and Technology (ICCET), 2nd International Conference on**. China: IEEE, 2010. v. 5, p. 619–623.

BELAN, P.A.; ARAUJO, A. S.; LIBRANTZ, A. F. H. Segmentation-free approaches of computer vision for automatic calibration of digital and analog instruments. **Measurement**, v. 46, p. 177–184, 2013.

CASTELLS-RUFAS, David; CARRABINA, Jordi. Camera-based digit recognition system. In: **Electronics Circuits and Systems, 13th IEEE International Conference on**. Nice, France: IEEE, 2006. p. 756–759.

EDWARD, Cephas P. Support vector machine based automatic electric meter reading system. In: **Computational Intelligence and Computing Research (ICCIC), IEEE International Conference on**. India: IEEE, 2013. p. 1–5.

GAZCÓN, Nicolás F.; CHESÑEVAR, Carlos I.; CASTRO, Silvia M. Automatic vehicle identification for argentinean license plates using intelligent template matching. **Pattern Recognition Letters**, v. 33, n. 9, p. 1066–1074, 2012.

HUANG, Guang-Bin; ZHU, Qin-Yu; SIEW, Chee-Kheong. Extreme learning machine: theory and applications. **Neurocomputing**, v. 70, n. 1, p. 489–501, 2007.

LEI, Haijun et al. Numeral recognition of power meter on a handheld terminal. In: **Proceedings of the Third International Symposium on Electronic Commerce and Security Workshops**. China: [s.n.], 2010. p. 76–79.

NAVA-ORTIZ, M.; GÓMEZ, Wilfrido; DIAZ-PEREZ, A. Digit recognition system for camera mobile phones. In: **Electrical Engineering Computing Science and Automatic Control (CCE), 8th International Conference on**. México: [s.n.], 2011. p. 1–5.

NODARI, Angelo; GALLO, Ignazio. A multi-neural network approach to image detection and segmentation of gas meter counter. In: **12th IAPR Conference on Machine Vision Applications**. Varese, Itália: [s.n.], 2011. p. 239–242.

OBINATA, Goro; DUTTA, Ashish. **Vision Systems: Segmentation and Pattern Recognition**. [S.l.]: I-Tech Education and Publishing, 2007.

OLIVEIRA, Danilo M.; CRUZ, Ruironald S.; BENSEBAA, Kamel. Automatic numeric characters recognition of kilowatt-hour meter. In: **Signal-Image Technology & Internet-Based Systems (SITIS), Fifth International Conference on**. Morocco: IEEE, 2009. p. 107–111.

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

PITAS, Ioannis. **Digital Image Processing Algorithms and Applications**. USA: John Wiley and Sons, 2000.

RUDEK, Marcelo. **Uma proposta para indexação e recuperação automática de imagens e reconhecimento de cheques bancários baseadas no vetor de comportamento**. 1999. Dissertação (Mestrado) — CPGEI CEFET-PR, Curitiba.

SANTOS, Eulanda M. dos. **Teoria e aplicação de support vector machines à aprendizagem e reconhecimento de objetos baseado na aparência**. 2002. Tese (Doutorado) — Universidade Federal da Paraíba.

SHU, Dongmei; MA, Shushua; JING, Chunguo. Study of the automatic reading of watt meter based on image processing technology. In: **2nd IEEE Conference on Industrial Electronics and Applications**. China: IEEE, 2007. p. 2214–2217.

SILVESTRE, Leonardo J. **Regularização de Extreme Learning Machines: uma abordagem com matrizes de afinidade**. 2015. Tese (Doutorado) — UFMG, Belo Horizonte.

SMITH, Ray. An overview of the tesseract ocr engine. In: IEEE. **Proceedings of Document analysis and Recognition, Ninth International Conference**. Brazil: ICDAR, 2007.

STONE, Mervyn. Cross-validators choice and assessment of statistical predictions. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 111–147, 1974.

VANETTI, Marco; GALLO, Ignazio; NODARI, Angelo. Gas meter reading from real world images using a multi-net system. **Pattern Recognition Letters**, v. 34, n. 5, p. 519–546, 2012.

VAPNIK, Vladimir. **The nature of statistical learning theory**. USA: Springer Science & Business Media, 2013.

WEN, Ying et al. An algorithm for license plate recognition applied to intelligent transportation system. **IEEE Transactions on Intelligent Transportation Systems**, v. 12, n. 3, p. 830–845, 2011.

ZHANG, Jing et al. A robust approach for recognition of text embedded in natural scenes. In: **Pattern Recognition, 16th International Conference on**. Canada: IEEE, 2002. p. 204–207.

ZHAO, Limeng et al. Research of digital meter identifier based on dsp and neural network. In: **Imaging Systems and Techniques, IEEE International Workshop on**. China: IEEE, 2009. p. 402–406.

ZHAO, Wei; GAO, MingYu; HE, Zhi Wei. Character recognition based on improved bp neural network. In: **Instrumentation, Measurement, Computer, Communication and Control (IMCCC), Third International Conference on**. [S.l.]: IEEE. p. 754–757.