



Apolo Takeshi Arai Batista

**Evolução no Processamento de Imagens
usando Redes Neurais**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Especialista pelo Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio.

Orientador: Prof. Leonardo Alfredo Forero Mendoza DSc

Rio de Janeiro
Abril de 2018



Apolo Takeshi Arai Batista

Evolução no Processamento de Imagens usando Redes Neurais

Dissertação apresentada como requisito parcial para obtenção do grau de Especialista pelo Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Leonardo Alfredo Forero Mendoza DSc
Orientador
Departamento de Engenharia Elétrica – PUC-Rio

Rio de Janeiro, 19 de Abril de 2018

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Apolo Takeshi Arai Batista

Graduou-se em Engenharia de Produção pela Universidade Federal do Rio de Janeiro (UFRJ, Brasil). Especializou-se em programas de qualidade obtendo a certificação Black Belt, dando continuidade aos conhecimentos obtidos em sua vivência junto à GE - Oil and Gas. Em seguida buscou aprofundar-se na parte de processamento de dados e visão computacional encontrando o Departamento de Engenharia Elétrica da PUC-Rio, especializando-se na área de visão computacional, processamento e análise digital de imagens. Atualmente cursa o Mestrado no Instituto Militar de Engenharia (IME, Brasil).

Ficha Catalográfica

Arai Batista, Apolo Takeshi

Evolução no Processamento de Imagens usando Redes Neurais / Apolo Takeshi Arai Batista; orientador: Leonardo Alfredo Forero Mendoza DSc. – Rio de Janeiro: PUC-Rio, Departamento de Engenharia Elétrica, 2018.

v., 31 f: il. color. ; 30 cm

Dissertação (Especialista) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Engenharia Elétrica – Teses. 3. Visão Computacional;. 4. Redes Neurais;. 5. Survey;. 6. Análise de Imagens;. 7. Classificação;. 8. Processamento de Imagens.. I. Mendoza, Leonardo Alfredo Forero. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 620.11

Agradecimentos

Gostaria de agradecer a todos professores do curso de Business Intelligence da PUC que se dispuseram para transferir conhecimentos à elaboração deste trabalho. Em especial ao meu orientador Leonardo Mendoza que ficou por horas a fio comigo num chat e ao Cristian Muñoz que me salvou várias vezes na parte de código, Jupyter e recuperação quando o servidor com toda minha pesquisa caiu.

Gostaria de agradecer muito à minha turma do BIMaster 2016, principalmente à galera do busão/metrô e à minha parceira de Startup! Não vou por seu nome aqui porque sei que você não gosta de ser facilmente localizável!

Obrigado a todos que compuseram o curso, desde às pessoas na secretaria que resolveram meus problemas com boletos, os técnicos que conseguiram manter as aulas presenciais e à distância funcionando em pleno vapor e à PUC por oferecer essa infraestrutura sem a qual este aprendizado não seria possível.

Resumo

Arai Batista, Apolo Takeshi; Mendoza, Leonardo Alfredo Forero. **Evolução no Processamento de Imagens usando Redes Neurais**. Rio de Janeiro, 2018. 31p. Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A visão computacional tem sido um debatida extensivamente nos últimos anos. Isto, aliado a um crescente avanço computacional em hardware (placas gráficas), novos algoritmos e formas de processar os dados, permite o surgimento de novas aplicações e questionamentos sobre o assunto. Este trabalho visa consolidar a informação no quesito reconhecimento de imagens de modo a servir como fonte para pesquisas futuras para processamento de imagens. Foram feitos dois experimentos, um com a base MNIST e outro com o reconhecimento facial. A nível de reconhecimento de imagens, o rosto humano seria qualificado como um objeto dinâmico e complexo. Os resultados dos experimentos realizados evidenciam que as novas formas de utilização das redes neurais as tornaram mais eficientes no reconhecimento de imagens. Os critérios propostos para o experimento e os resultados destes serão discutidos detalhadamente.

Palavras-chave

Visão Computacional; Redes Neurais; Survey; Análise de Imagens; Classificação; Processamento de Imagens.

Abstract

Arai Batista, Apolo Takeshi; Mendoza, Leonardo Alfredo Forero (Advisor). **Neural Net processing images evolution**. Rio de Janeiro, 2018. 31p. Dissertação de mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Computer Vision has received much attention in the last few years. The recently breakthrough of technology (more powerful graphics card and new algorithm) has made companies been more creative regarding their solution to save time and reduce costs in their applications. This research has been built in order to fulfill their needs of knowing the state of art regarding image processing. The experimental procedure (MNIST and facial recognition due its complexity and dynamics) and results from the experiments performed are discussed in detail.

Keywords

Computacional view; Neural Net; Survey; Image Analysis; Classification; Image processing.

Sumário

1	Introdução	5
2	Revisão Teórica Redes Neurais e Processamento de Imagens	7
2.1	MLP	9
2.2	SGD	9
2.3	Dropout	11
2.4	Rmsprop	11
2.5	Redes Convolucionais	11
3	Conceitos Básicos em Processamento de Imagens para Reconhecimento Facial	14
4	Discussões	17
5	Aplicações Práticas	18
6	Experimento	20
7	Conclusão e Trabalhos Futuros	22
8	Referencias Bibliográficas:	24

Lista de figuras

Figura 2.1	Neurônio ArtificialBarreto (2002)	10
Figura 2.2	Exemplos de funções de ativação. Adaptado de Faceli (2011)	10
Figura 2.3	Timeline da Evolução ao longo do tempo – Redes Neurais	10
Figura 2.4	Arquitetura de Rede Multi Layer Perceptron (MLP)BONIFACIO (2010)	13
Figura 2.5	Comportamento com elevada taxa de aprendizadoHinton et al. (2012)	13
Figura 2.6	Rede Convolutional para processamento de imagem LeCun e Bengio (1995)	13
Figura 3.1	Exemplo de captura de imagens com sujeito a ser re-cohecido	14
Figura 3.2	Etapas desde a captura da imagem até o reconhecimento de indivíduo	15

Lista de tabelas

Tabela 2.1	Evolução ao longo do tempo – Redes Neurais	8
Tabela 2.2	Vantagens e desvantagens – Redes Neurais	10
Tabela 3.1	Principais problemas por etapa	15
Tabela 3.2	Objetivos possíveis ao escolher algoritmos	16
Tabela 4.1	Redes Neurais Convolucionais	17
Tabela 6.1	Experimentos Realizados Base MNIST	21
Tabela 6.2	Experimentos Realizados Base YaleFaces	21
Tabela 7.1	Foco e inspiração para trabalhos futuros	22

1

Introdução

A visão computacional objetiva replicar a visão humana (HJELMÅS; LOW, 2001). Com o avanço tecnológico, surgimento de novos algoritmos e placas gráficas, cada vez mais as empresas têm sido mais criativas em suas soluções para economizar tempo e reduzir custos de suas aplicações. O reconhecimento de imagens permite análises mais assertivas para identificação de um indivíduo (impressões digitais Meva et al. (2012), reconhecimento de caligrafia Bortolozzi et al. (2005) e reconhecimento facial Lin (2000)) ou para identificar se uma fotografia é real ou forjada Lukáš et al. (2006), por exemplo.

A tarefa do reconhecimento de faces pode ser simplificada como analisar e detectar um número de faces numa determinada imagem apresentada. Que passaria pelos passos de fragmentar a imagem, extrair os dados relevantes separando as faces do fundo e verificar individualmente as faces encontradas e, idealmente, tratar questões como distorções que incorram devido a câmera utilizada, distância focal, balanceamento de luminosidade e correção de cores. Embora existam trabalhos publicados desde 1965, ainda hoje o reconhecimento de faces é um tema de elevada complexidade. Muitas aplicações se baseiam no reconhecimento facial como reconhecimento facial e verificação, vigilância por meio do monitoramento de faces, análise do comportamento facial, reconhecimento de atributos faciais (obtenção de gênero e idade, "entendimento do belo").

Este trabalho visa consolidar a informação no quesito reconhecimento de imagens, com ênfase na evolução de métodos de análises. De modo a servir como fonte para pesquisas futuras. É feito o levantamento teórico e apresentação das principais bases de dados utilizadas para validação de modelos. Ao analisar a abordagem teórica, pode-se perceber que a frequência e a velocidade com que novos algoritmos e novas técnicas, vêm sendo desenvolvidos está cada vez menor.

Este texto contém outras 7 sessões. Na seção 2, será feita uma revisão teórica detalhada sobre as redes neurais, desde seu primeiro surgimento até técnicas mais recentes. Na seção 3, serão vistos os processamentos de imagens para reconhecimento facial, desde seu diagrama simplificado até os principais objetivos de sua utilização. Na seção 4 são feitas algumas discussões acerca da

tecnologia que está disponível atualmente e a apresentação da escolha que será feita no experimento. Ao passo que na seção 5 é apresentada sua utilização atual. Por fim, será feita a conclusão além de citar as principais inspirações para investigações futuras, bem como problemas que possam se encontrar em aberto.

A técnica de redes neurais se inicia antes mesmo de estarmos falando de computadores, mas do corpo humano. Com a descoberta da possibilidade de se isolar neurônios e estudar seu funcionamento, que só aconteceu em 1875 por um italiano chamado Camillo Golgi (BARRETO, 2002 apud HUBEL, 1979), permitiu-se conhecer melhor o organismo humano e se traçar paralelos, questionar e entender sua funcionalidade.

Esse entendimento fez com que pudessem ser propostos modelos para o funcionamento de neurônios, a partir dos quais estudos foram formalizados em teoremas pela primeira vez por McCulloch e Pitts (1943). Baseados nos neurônios humanos foram estruturadas equações matemáticas para resolver problemas lineares e essa forma de se ajustar a informação foi concebida como neurônio artificial. Tais neurônios eram capazes de operar independentemente uns dos outros, conforme (FACELI, 2011, p.108) explica: “os neurônios executavam funções lógicas simples e cada um podia executar uma função diferente”.

Na concepção do McCulloch e Pitts (1943) o neurônio utilizava funções de ativações lineares, que seguiam um fluxo simétrico e constante para seu comportamento. Na Figura 2.1, conforme levantado por Barreto (2002) é exemplificado um neurônio composto por entradas w e u que são combinadas por meio de uma função f . Essa junção provoca uma reação que vai caminhar a produzir um estado de ativação do neurônio que ocorrerá por meio da função explícita em g que irá produzir a saída correspondente ao neurônio. Por muitas vezes existe um valor auxiliar chamado θ que é utilizado para representar a polarização, valor a partir do qual a saída se torna nula.

No que tange as funções de ativações, Faceli (2011) apresenta 3 funções, conforme apresentado na Figura 2.2, ressaltando que a função Sigmoidal permite diferentes inclinações, pois representa uma aproximação contínua e diferenciável da função limiar.

Para facilitar a contextualização histórica dos envolvidos em grande parte da evolução de redes neurais, as informações foram consolidadas na Tabela 2.1 e na linha do tempo conforme Figura 2.3.

Tendo entendido como um neurônio (muitas vezes também chamado de perceptron) funciona, é necessário destacar Rumelhart et al. (1986), que

Tabela 2.1: Evolução ao longo do tempo – Redes Neurais

Ocorrido	Autor e Ano
Primeiro Modelo de Neurônio Artificial - Perceptron	(MCCULLOCH; PITTS, 1943)
Cibernética: estudo do controle e comunicação entre animal e máquina	(WIENER, 1948)
Desenvolveu a conexão de pesos variáveis, parte fundamental da regra de aprendizado nas redes neurais	(HEBB, 1950)
Invenção do algoritmo Perceptron, discriminante linear cujos pesos permanecem imutáveis no caso de uma classificação bem feita dos dados	(ROSENBLATT, 1957)
Publicação de um livro chamado 'perceptron'. Onde eles provaram que embora os perceptrons pudessem ser apresentados a aprender qualquer conteúdo cuja representação fosse possível eles iriam representar muito pouco, por exemplo. Um perceptron de duas entradas não poderia ser treinado para reconhecer quando duas entradas fossem diferentes. Além da disputa entre portas lógicas e neurônios artificiais para determinar unidade básica de processamento	(MINSKY; PAPERT, 1969)
Mapas Auto-Organizáveis com o objetivo de reproduzir computacionalmente experimentos com estímulos visuais.	(MALSBURG, 1973)
A criação de uma função de energia, que, trabalhando com pesos simétricos assegura uma estabilidade na rede. Que pode ser exemplificado com a decomposição de algo mais complexo num vetor para depois sua recomposição.	(HOPFIELD, 1982)
Mapas Auto-Organizáveis cujo modelo baseia-se na procura do neurônio cujos pesos são mais próximos de um determinado padrão de entrada (com a menor distância Euclidiana) e no aumento da similaridade entre eles (padrão de entrada e pesos do neurônio vencedor)	(KOHONEN, 1982)
Desenvolvimento de um sistema de aprendizagem, que consiste em usar dois neurônios diferentes com elementos adaptativos num Elemento de Busca Associativa (tradução livre de Associative Search Element - ASE) que permite aplicar o aprendizado por reforço às redes neurais, implementando e ajustando a regra de controle ou política de decisão sobre o que prossegue na rede.	(BARTO et al., 1983)
Aprofundando os estudos iniciados por Barto, desenvolveu o Elemento Adaptativo crítico, tradução livre de (Adaptive Critic Element - ACE) que utilizava a falha do sinal para aprender como gerar informação útil a todo momento para avaliar o controle da decisão.	(SUTTON, 1984)
Backpropagation - método mais popular de treinamento de perceptrons de múltiplas camadas	(RUMELHART et al., 1985)

permitiram que os neurônios pudessem ir para outro nível. Uma vez que eles desenvolveram o algoritmo back propagation, fizeram com que neurônios pudessem ser organizados com o intuito de formar uma Rede Neural. Nesta, os neurônios estão organizados ou dinamicamente ou estaticamente, com suas funções de ativações explícitas e formas de aprendizado na rede e perturbações que serão causadas a esta rede. O formato mais simples de redes neurais são as chamadas Redes Neurais Feedforward ou Feed forward Neural Networks (FFNN), onde todos os neurônios de uma camada são conectados com todos os neurônios da camada seguinte, possuindo dois algoritmos possíveis: forward propagation e back propagation, que são usados para computar o vetor de saída de ativação da rede e reajustar os parâmetros da rede com base em alguma medida de erro Rocha (2015). Desse modo, embora inspirados na fisiologia e biologia, a rede neural é muitas vezes considerada tão capaz quanto os neurônios humanos, devido a sua capacidade de identificar padrões, porém, conforme a Tabela 2.2.

2.1

MLP

A rede neural, também pode ser referida como MLP (Multi Layer Perceptron), pois corresponde a mais de um perceptron estruturado seguidamente. Com ela foi possível que se aproximasse qualquer função contínua num compacto conjunto com uma acurácia desde que sejam providos neurônios suficientes na camada escondida Lawrence et al. (1996). Para facilitar o que seria uma camada escondida e qual a diferença entre o perceptron que continha apenas camadas de entrada e de saída, veja a figura a seguir, quaisquer camadas que não sejam a camada de entrada e a camada de saída serão chamadas de camadas ocultas ou hidden layers Rocha (2015):

2.2

SGD

A performance do Stochastic Gradient Descent (SGD) ou Gradiente Estocástico Descendente, à tradução livre. Depende criticamente de como estão sendo manipuladas as taxas de aprendizado ao longo do tempo Schaul et al. (2012). É uma técnica focada no aprendizado da rede neural. Que visa mover o ponto para que, somado com o “momento” ele não fique parado num mínimo local e isso pode dar errado quando o peso é tão grande que faz com que haja uma fuga do mínimo local, com os valores convergindo para longe do mínimo quando era desejado justamente o oposto Hinton et al. (2012), conforme mostrado na figura abaixo quando a taxa de aprendizado é muito elevada.

Desse modo percebe-se que a o gradiente deve ser estruturado para causar

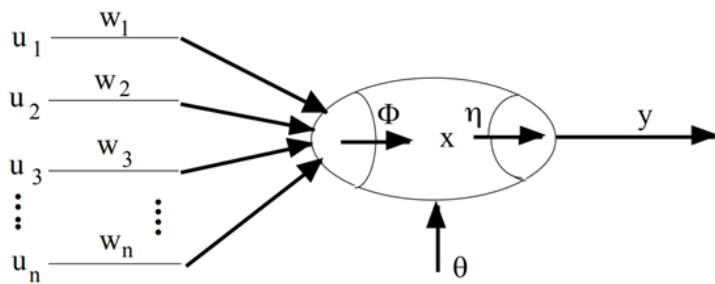


Figura 2.1: Neurônio ArtificialBarreto (2002)

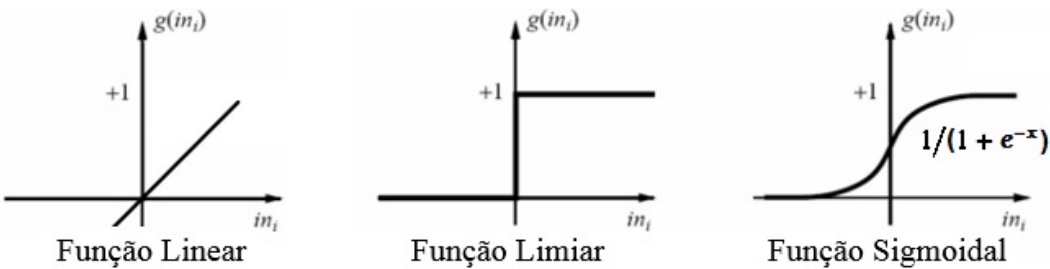


Figura 2.2: Exemplos de funções de ativação. Adaptado de Faceli (2011)

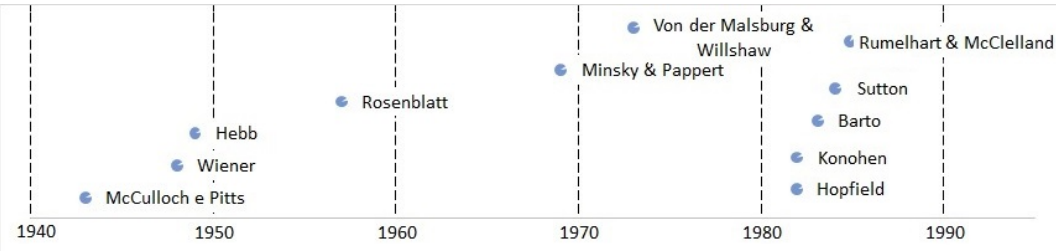


Figura 2.3: Timeline da Evolução ao longo do tempo – Redes Neurais

Tabela 2.2: Vantagens e desvantagens – Redes Neurais

Vantagens	Desvantagens
Capaz de prosseguir ainda que um elemento falhe, natureza de processamento em paralelo	A arquitetura de uma rede neural é diferente da arquitetura de microprocessadores ainda que pareça copiá-la
Pós treinamento e pesos, não necessita ser reprogramada. Capaz de realizar tarefas que uma linear não poderia.	Requer treinamento: muito custoso computacionalmente.
[tl]Pode ser implementada: 1. em qualquer aplicação 2. sem grandes problemas	Requer tempo elevado de processamento para uma grande rede neural

uma perturbação na sua curva de aprendizado de modo a dificultar que fique preso num mínimo local quanto não seja direcionado para longe de todos os mínimos de modo a nunca apresentar um resultado.

2.3

Dropout

O dropout é uma técnica de regularização Rocha (2015); Hinton et al. (2012) pois permite que haja a desconexão na saída dos neurônios de uma rede neural profunda. Seu funcionamento se dá por meio da desativação de alguns neurônios da camada, de modo que haja um treinamento utilizando apenas parte dos neurônios a cada treinamento. Com isso a rede fica com diferentes neurônios treinados para a resolução do mesmo problema de modo mais eficiente, pois ela se torna mais genérica e capaz de fazer com menos neurônios o mesmo serviço. Assim sendo, ela é capaz de quando, após o treinamento, for processar a informação, poder ter mais neurônios que vão direcioná-la no caminho certo.

2.4

Rmsprop

Conforme Dauphin et al. (2015) RMSProp é um método de taxa de aprendizado adaptativo que encontrou muito sucesso na prática (BENGIO, 2015 apud TIELEMAN HINTON, 2012, KORJUS et ali; CARLSON et ali, 2015). A funcionalidade era dada devido a normalização dos gradientes por uma exponencial movendo a média da magnitude do gradiente para cada parâmetro. Então Dauphin et al. (2015) mostra que o RMSProp é, de fato, um estimador baseado na pre-condicional de equilíbrio. O RMSProp surge devido à necessidade de se acreditar que não se deve trabalhar com a mesma taxa de aprendizado ao longo de todos os pesos da rede neural e que esta deve poder mudar durante o aprendizado de modo que seria pouco interessante ter uma única taxa de aprendizado para toda a rede neural Hinton et al. (2012).

2.5

Redes Convolucionais

Fukushima (1980) desenvolveu um modelo não supervisionado chamado Neocognitron que consistia em observar uma imagem, varre-la da esquerda pra direita, de cima para baixo, linha a linha, codificando ela e então descodificando e interpretando seus estímulos. Esse modelo serviu como base para que pudesse ser feito a codificação de caracteres escritos à mão livre e sua interpretação

LeCun e Bengio (1995). Conforme a figura abaixo é possível visualizar essas etapas:

Por meio das redes convolucionais se tornou possível fazer o processamento de imagens sem serem necessários tratamentos prévios antes de inserí-la numa rede neural, uma vez que da imagem são extraídas todas as características que servirão de input para um método que irá avaliá-lo. Então resumidamente, ganhou-se muita possibilidade de processamento de objetos mais complexos. Não apenas isso como a evolução tecnológica permitiu a utilização de computação paralela onde num exemplo desses poderíamos ter diversos trechos do processamento sendo feito em uma máquina diferente. Na Figura 2.6 estamos vendo o input sendo a imagem inicial. Seguidamente essa imagem se transforma em 4 imagens diferentes, podendo por exemplo iniciar sua varredura em posições diferentes ex.: (0,0); (4,0); (0,4) e (4,4); gerando assim, 4 mapas de features. Em seguida cada um desses mapas é reduzido pela metade segundo um algoritmo de identificação de símbolos, e novamente são feitas outras varreduras começando em outros pontos. E esse processo se repete até que se tenha um vetor que represente qualitativamente a imagem.

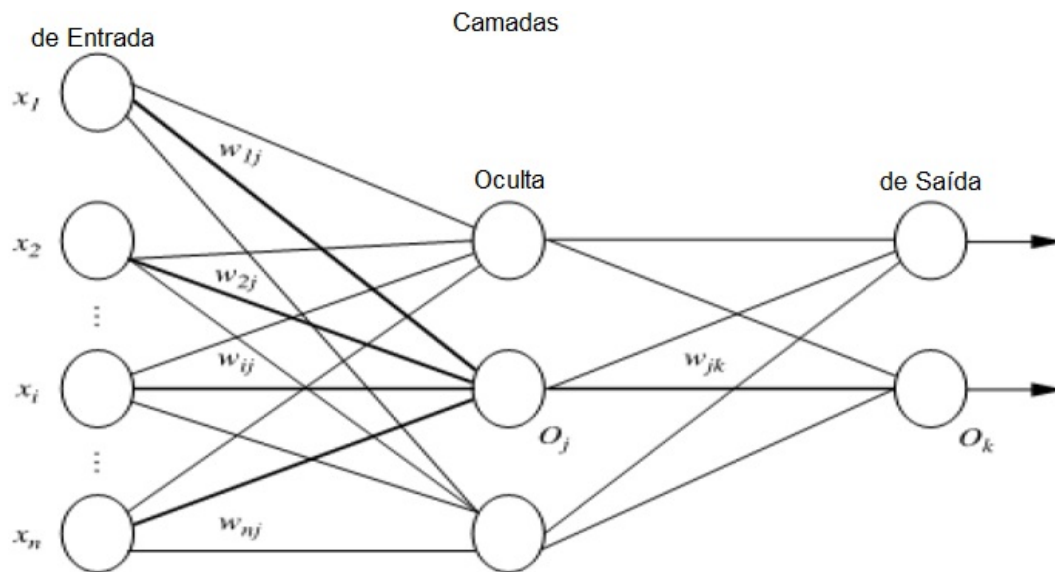


Figura 2.4: Arquitetura de Rede Multi Layer Perceptron (MLP) BONIFACIO (2010)

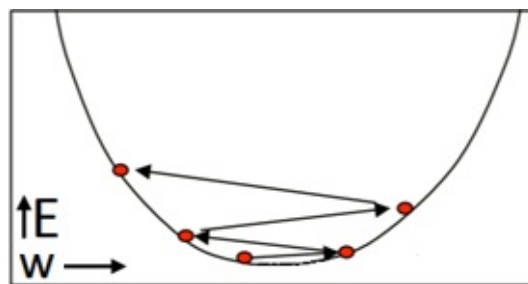


Figura 2.5: Comportamento com elevada taxa de aprendizado Hinton et al. (2012)

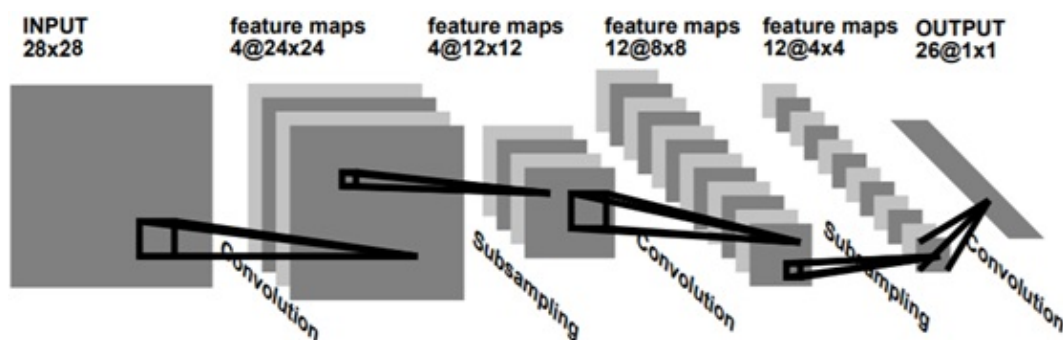


Figura 2.6: Rede Convolucional para processamento de imagem LeCun e Bengio (1995)

3

Conceitos Básicos em Processamento de Imagens para Reconhecimento Facial

Visando direcionar a discussão do atual Estado da Arte da visão computacional em relação ao reconhecimento facial, é feito um esquema simples representado na Figura 3.1.

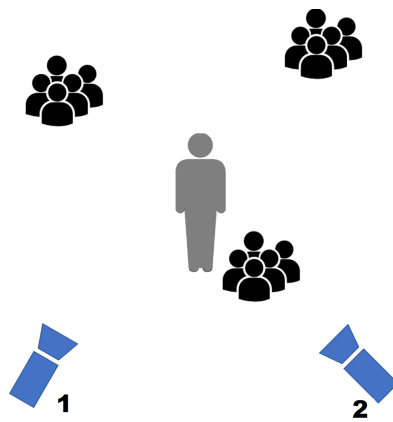


Figura 3.1: Exemplo de captura de imagens com sujeito a ser reconhecido

Nele temos um cenário composto por duas câmeras (1 e 2) e diversos indivíduos à diferentes distâncias da mesma. Em cinza, o indivíduo que queremos identificar.

Em seguida, na Figura 3.2 é apresentado um esboço simples de como o processo se dá desde a captura dos dados simples, no caso da imagem capturada por cada câmera e sua saída com o reconhecimento do indivíduo.

Na Tabela 3.1 são revistas as etapas apresentadas pela Figura 3.2 e expostas as principais questões acerca de cada uma das etapas apresentadas.

Já na Etapa Rodar Algoritmo, é percebido que os algoritmos poderão variar de acordo com o objetivo do que se deseja obter com cada um deles, podendo ser usados separadamente ou em conjuntos, como podem ser observados na Tabela 3.2

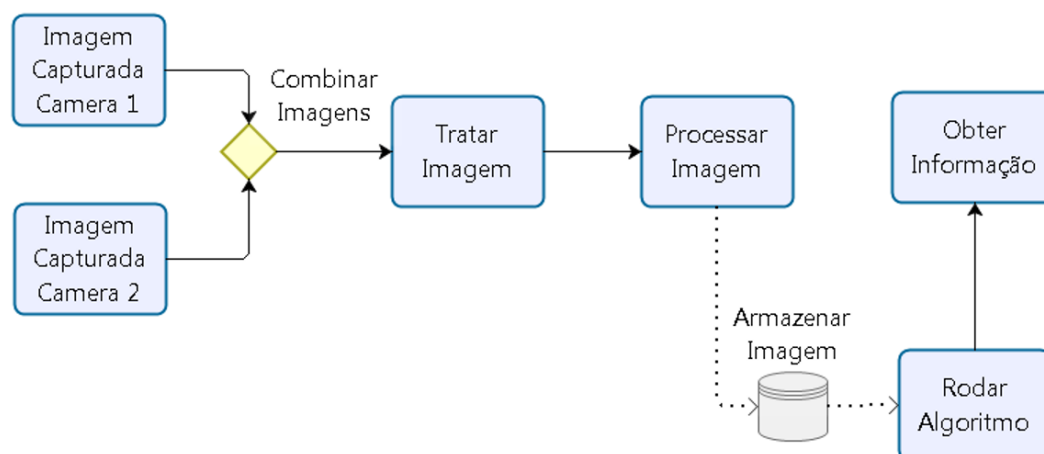


Figura 3.2: Etapas desde a captura da imagem até o reconhecimento de indivíduo

Tabela 3.1: Principais problemas por etapa

Etapa	Problema	Fonte
Imagem Capturada	Baixa Resolução das Câmeras	(ZAFEIRIOU et al., 2015)
Combinar Imagem	Diferença de Cores	(WANG et al., 2015)
	Diferença de Luminosidade	(QIAN et al., 2017)
	Diferença de distorção	(KUME et al., 2015)
	Diferença de distorção	(PENG et al., 2017)
Tratar Imagem	Elevar a qualidade da imagem	(NIE et al., 2017)
	Tratar de foco	(MOSLEH et al., 2018)
	Correção de Cores	(RAMA VARIOR et al., 2016)
	Correção de Cores	(LUKAC; PLATANIOTIS, 2005)
	Correção de Cores	(JANG et al., 2017)
	Inserção de Filtros	(ZHU et al., 2015)
	Inserção de Filtros	(LI; ZHENG, 2018)
Processar Imagem	Descrição de Cenário	(GONFAUS et al., 2015)
	Descrição de Cenário	(TUNG; LITTLE, 2015)
	Mapping da informação	(CHENG et al., 2017)
	Mapping da informação	(NIU et al., 2017)
	Mapping da informação	(MILLER et al., 2005)
	Mapping da informação	(WANG et al., 2015)
Armazenar Imagem	Risco de ataques	(ZHANG et al., 2017)
	Risco de ataques	(TIRUNAGARI et al., 2015)
	Risco de ataques	(H. FARID, 2009)

Tabela 3.2: Objetivos possíveis ao escolher algoritmos

Etapa	Detalhamento	Fonte
Reconhecimento Facial e Verificação		(SCHROFF; PHILBIN, 2015) (ZHAO et al., 2003) (CHELLAPPA et al., 1995) (SAMAL; IYENGAR, 1992) (VIOLA; JONES, 2001)
Vigilância por meio do monitoramento de Faces		(KALAL et al., 2010)
Análise do comportamento Facial		(PANTIE; ROTH-KRANTZ, 2000)
Reconhecimento de Atributos Faciais	Genero e Idade "Embelezamento"	(NAYAR, 2009) (FU et al., 2010) (LAURENTINI; BOTTINO, 2014)
Transformação e Re-iluminação Facial		(WANG et al., 2007)
Reconstrução de Formatos Faciais		(BLANZ; VETTER, 1999)
Recuperação de Imagem e Video		(KEMELMACHER-SHLIZERMAN et al., 2011)
Estilização de Rostos (filtros)	Desenho de Contorno ou linhas Rascunhado Caricaturização Cartunização	(CHEN et al., 2002) (WANG et al., 2013) (TSENG; LIEN, 2012) (ZHANG et al., 2017)
Comportamento anômalo dentro de multidão		(YUAN et al., 2015)

4

Discussões

Alguns trabalhos apresentam fragmentos similares como (SCHROFF; PHILBIN, 2015; WANG et al., 2014; ZHU et al., 2014). Onde Schroff e Philbin (2015) propõe um método que aprende o espaço da distância Euclidiana automaticamente, enquanto que Wang et al. (2014) propõe o ranqueamento de imagens por significado semântico e similaridade visual e enquanto Zhu et al. (2014) redimensiona o rosto trazendo-o para um formato visual "canônico", ou seja, frontal e sem expressão, Taigman et al. (2014) se limita a realizar a transformação do rosto para um modelo 3d do mesmo.

As técnicas mais recentes fazem uso do reconhecimento de imagens utilizando redes profundas, geralmente com foco em aprender baseada nos pixels do rosto, limitada à utilização de um grande banco de dados com rostos identificados visando obter aprendizados como variações de luz, posicoes do rosto ou influencia de plano de fundo. Parece válida a utilização das redes profundas apresentadas na tabela a seguir, haja visto sua extensa utilização nos trabalhos recentes, claro que por meio de algumas variações como inclusão ou exclusão de camadas. Aproveitando a utilização do ImageNet, campeão da competição em 2014, e o ConvNet.

Tabela 4.1: Redes Neurais Convolucionais

Rede Neural	Fonte
ConvNet	(LECUN et al., 1989) (RUMELHART et al., 1986) (SUN et al., 2015)
Imagenet 2014	(SZEGEDY et al., 2014)

5

Aplicações Práticas

Tendo sido melhor explicado o funcionamento do algoritmo, é possível perceber que, se comparado com um funcionário que faça uma operação repetitiva, a rede neural é mais interessante. Diferentemente de uma pessoa a rede neural não se cansaria nem ficaria fatigada, não requerendo tempo para ir ao banheiro ou almoçar. Com isso diversos processadores e sensores atuais são construídos com algoritmos de redes neurais que ao invés de apresentar todo o cenário a um interprete já “filtra” e direciona a ele a informação destacada (QIN, 1997). Embora cada vez mais estejam sendo utilizadas redes neurais, seu detalhamento exato não é muito explícito. Como explicado na seção anterior o neurônio lida apenas com as mais básicas informações e detecta padrões entre suas relações, gerando assim uma saída que será apresentada. Porém o que levou o neurônio a atuar daquela forma não é explícito. Muitas vezes sendo de grande dificuldade explicar o racional de um resultado obtido, podendo apenas inferir a relação entre os resultados exibidos com o conhecimento de um especialista. Essa habilidade, que poderia ser vista como uma desvantagem, se bem usada pode se tornar uma boa aliada. Como por exemplo ao controlar múltiplos sistemas pois, como os algoritmos aprendem com exemplos, no caso de aprendizado supervisionado, tem boas habilidades de generalização e conseguem performar eficientemente em tempo real (MOODY, 1989). Por exemplo, na indústria de óleo e gás, poderíamos usar redes neurais para lidar com a destilação (FADHIL AHMED; KHALAF, 2015). Enquanto que como apoiadora à processos de engenharia (civil (FLOOD et ali, 1994), ambiental (SHANIN, 2001)). Somado a isto, aproveitando suas habilidades em reconhecer padrões (BENGIO, 2015b) e a habilidade de classificar informações seria possível de utilizá-la dentro da medicina (AL-SHAYEA, 2011), bancos com suas aprovações de crédito (WU; HONG, 2005), mercado financeiro com a previsão de ações (GHAZNAVI et al., 2016). Bem como o reconhecimento anômalo de padrões em meio a multidões, detectando movimentos suspeitos de indivíduos que não deveriam transitar por aquelas áreas (VATSARAJ, 2017). Outra aplicação que tem se destacado é em marketing, onde a rede neural pode ser utilizada para fazer segmentação de mercado de acordo com os interesses dos clientes (PIATETSKY-SHAPIRO,

1996), retenção de clientes (FERREIRA, 2005). A aplicação de rede neural, se for entrado em redes convolucionais ou profundas, pode-se destacar inclusive a capacidade de aprender e lidar com a identificação de emoções na fala (HAN et al., 2014), nos textos (DOS SANTOS et ali, 2014) ou até mesmo na identificação de padrões musicais (ULLRICH, 2014). Tendo visto o amplo escopo de aplicação, este levantamento visa levantar técnicas de aplicação para o processamento de imagens usando redes neurais. Embora seja possível se pensar em processamento de vídeo, onde o vídeo seria uma sucessão de frames, e assim, de imagens, o processamento de vídeos não corresponde ao escopo deste trabalho. Para auxiliar a visualização da evolução temporal das redes neurais, estruturou-se um experimento. Este, de fácil replicação, visa ilustrar a mudança do algoritmo ao longo dos anos, que embora tenha ganho mais complexidade também ganhou mais eficiência.

6 Experimento

Para este estudo foram utilizados os seguintes algoritmos:

1. Perceptron + SGD
2. MLP + SGD
3. MLP + SGD + Dropout
4. MLP + Dropout + Rmsprop
5. ConvNet - LeNet

Todos foram testados usando os mesmos dataset e mesmos seeds para reprodutibilidade. Variando sempre o número de épocas a que serão submetidos, sendo elas, 20, 50, 100 e 200. Em todas as configurações foram utilizados 128 neurônios nas camadas escondidas. Todos os resultados foram coletados após 10 rodadas de experimentos, variando os seeds entre 1661 e 1671. São apresentados os resultados médios obtidos por cada algoritmo a cada intervalo igual de épocas. Podemos observar que com a Base MNIST (que é uma base de imagens desenhadas a mão que variam de 0 a 9) houve uma performance melhor que no dataset Yalefaces (composto por 13 imagens de 15 indivíduos) uma base de dados mais complexa. Uma exceção a esta observação é a rede Convolutional que, com suas múltiplas camadas, é capaz de observar e entender padrões complexos.

Tabela 6.1: Experimentos Realizados Base MNIST

Algoritmo	épocas	20	50	100	200
Perceptron + SGD	Test Score	0,3486	0,3071	0,2887	0,2780
	Test Accuracy	0,9059	0,9157	0,9205	0,9220
MLP + SGD	Test Score	0,1875	0,1171	0,0860	0,0781
	Test Accuracy	0,9460	0,9648	0,9741	0,9762
MLP + SGD + Dropout	Test Score	0,1971	0,1255	0,0933	0,0785
	Test Accuracy	0,9403	0,9610	0,9726	0,9774
MLP + Dropout + Rmsprop	Test Score	0,0986	0,1327	0,1655	0,1962
	Test Accuracy	0,9803	0,9771	0,9796	0,9765
ConvNet - LeNet	Test Score	0,0362	0,0343	0,0436	0,0437
	Test Accuracy	0,9921	0,9943	0,9940	0,9941

Tabela 6.2: Experimentos Realizados Base YaleFaces

Algoritmo	épocas	20	50	100	200
Perceptron + SGD	Test Score	14,6959	14,6959	14,6959	14,6959
	Test Accuracy	0,0882	0,0882	0,0882	0,0882
MLP + SGD	Test Score	15,1700	15,1700	15,1700	15,1700
	Test Accuracy	0,0588	0,0588	0,0588	0,0588
MLP + SGD + Dropout	Test Score	2,7092	2,7092	2,7092	2,7092
	Test Accuracy	0,0294	0,0294	0,0294	0,0294
MLP + Dropout + Rmsprop	Test Score	15,1700	15,1700	15,1700	15,6440
	Test Accuracy	0,0588	0,0588	0,0588	0,0294
ConvNet - LeNet	Test Score	1,0169	1,7799	1,4994	1,4886
	Test Accuracy	0,8529	0,7647	0,7647	0,7647

7

Conclusão e Trabalhos Futuros

Ao longo da pesquisa foram encontradas alguns focos comumente abordados sobre detecção de faces por múltiplas Visões e do aprendizado sem rótulos de subcategorias, onde os principais trabalhos são elucidados na Tabela 7.1.

Tabela 7.1: Foco e inspiração para trabalhos futuros

Foco	Problema	Fonte
Detecção de Faces por múltiplas visões	Cascata Paralelas	(BO WU et al., 2004)
	Estrutura Pirâmide	(LI et al., 2002)
	Árvores de Decisão	(JONES; VIOLA, 2003)
	Boost em Vetores Valorados	(CHANG HUANG et al., 2005)
Aprendizado sem rótulos de subcategorias	Boost em Vetores Valorados	(LIN; LIU, 2005)
	Clusterização e Treinamento	(SEEMANN et al., 2006)
	Aprendizado Baseado num exemplar	(SHAN et al., 2006)
	árvore de boost probabilístico	(TU, 2005)
	Clusterização com características selecionadas	(WU; NEVATIA, 2007)
	Classificador Múltiplo/Boost de categoria	(KIM; CIPOLLA, 2009)
	Classificador Múltiplo/Boost de categoria	(MICHAEL, 2012)
	Classificador Múltiplo/Boost de categoria	(ZHANG; ZHANG, 2009)

Foram observados alguns problemas em aberto como a melhoria da qualidade de uma imagem de baixa resolução capturada em um ambiente não controlado, bem como o reconhecimento de rostos em ambientes não controlados (comumente chamados na literatura como *'in the wild'*). Além disso, evidencia-se como um problema em aberto então a utilização de técnicas que já estão sendo empregadas no reconhecimento de objetos para verificação se melhoraria a acurácia do reconhecimento de rosto. Essa abordagem foi

inspirada por (HU et al., 2016) onde a lógica de ter poucas imagens num banco de dados original requeri um trabalho de modo a viabilizar a utilização de tal banco para reconhecimento de imagens, ignorando o fato da base ser extremamente limitada. Identificação de rostos em uma imagem seguida de sua projeção em 3D para estruturar um rosto (DING et al., 2015), utilização de algoritmo para melhorar os contornos e diminuir a ocorrência de falsos positivos, garantindo uma melhor acurácia no modelo de reconhecimento (AYTAR; ZISSERMAN, 2015) e, por fim, combinar (ZHANG et al., 2016) com (TABERNIK et al., 2015) visando gerar clusters que possam ser trabalhados como os apresentados por (WANG et al., 2015). Ficando indicado, assim, para trabalhos futuros a experimentação da utilização de técnicas semânticas para identificação de clusters de rostos em diferentes rostos, sobre as premissas de que por uma varredura num espaço menor de busca haveria ganhos de performance e acurácia no reconhecimento facial, uma vez que grande parte dos problemas são relativos a performance e num banco de dados muito grande, em cenários reais, a acurácia da rede neural torna-se menos elevada.

Como conclusão, podemos destacar que as aplicações com redes neurais nem sempre são 100

Como podemos observar com o experimento, as novas técnicas de redes neurais permitem que com menos ciclos (épocas) seja possível obter um resultado de acurácia mais elevado e a execução de mais épocas pouco interfere no resultado obtido. Mostrando que, nos dias de hoje, utilizar redes neurais pode ser uma boa alternativa pois com a computação paralela e o fato de que as redes neurais com mais camadas são capazes de ilustrar qualquer função matemática, quase qualquer problema poderá ser modelado utilizando redes neurais. Com isso, o domínio desta técnica se torna estratégico para todas as companhias.

Ou seja, caso uma empresa deseje entrar nesse meio, é necessário adquirir tal conhecimento e não apenas comprar um. Caso apenas se adquira, é bem provável que, em pouco tempo, a tecnologia esteja obsoleta. Isso faz com que, muitas vezes, possa levar companhias a terceirizar um serviço que será um diferencial estratégico no longo prazo. Então obter uma aplicação, que seja desenvolvida nessas bases sem obter seu racional é condenar-se a uma dependência do fornecedor no médio prazo. Então é necessário que sempre haja uma pesquisa para que se possa estar atualizado com as formas de utilização de redes neurais existentes.

Referencias Bibliográficas:

- AL-SHAYEA, Q. **Artificial Neural Networks in Medical Diagnosis**. ... Journal of Computer Science Issues (IJCSI), 8(2):150–154, 2011.
- AYTAR, Y.; ZISSERMAN, A. **Part level transfer regularization for enhancing exemplar SVMs**. Computer Vision and Image Understanding, 138:114–123, 2015.
- BARRETO, J. M. **Introdução às Redes Neurais Artificiais**. p. 57, 2002.
- BARTO, A. G.; SUTTON, R. S. ; ANDERSON, C. W. **Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems**. IEEE Transactions on Systems, Man and Cybernetics, SMC-13(5):834–846, 1983.
- BLANZ, V.; VETTER, T. **A morphable model for the synthesis of 3D faces**. Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99, p. 187–194, 1999.
- BO WU; HAIZHOU AI; CHANG HUANG ; SHIHONG LAO. **Fast rotation invariant multi-view face detection based on real adaboost**. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings., p. 79–84, 2004.
- BONIFACIO, F. N. **Comparação Entre As Redes Neurais Artificiais Mlp , Rbf**. PhD thesis, Universidade Estadual do Oeste do Paraná–Centro de Ciências Exatas e Tecnológicas-CCET, 2010.
- BORTOLOZZI, F.; JR, A. D. S. B. ; OLIVEIRA, L. **Recent advances in handwriting recognition**. Número January 2005. 2005.
- CHANG HUANG; HAIZHOU AI; YUAN LI ; SHIHONG LAO. **Vector boosting for rotation invariant multi-view face detection**. Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 1:446–453 Vol. 1, 2005.
- CHELLAPPA, R.; SIROHEY, S. ; WILSON, C. L. **Human and Machine Recognition of Faces: A Survey**. Proceedings of the IEEE, 83(5):705–741, 1995.

- CHEN, H.; ZHENG, N.-N.; LIANG, L.; LI, Y.; XU, Y.-Q. ; SHUM, H.-Y. **PicToon: A Personalized Image-based Cartoon System**. Proceedings of ACM International Conference on Multimedia, (49):171–178, 2002.
- CHENG, F.; HE, X. ; ZHANG, H. **Stacked Learning to Search for Scene Labeling**. IEEE Transactions on Image Processing, 26(4):1887–1898, 2017.
- DAUPHIN, Y. N.; DE VRIES, H.; CHUNG, J. ; BENGIO, Y. **Equilibrated adaptive learning rates for non-convex optimization**. 2015.
- DING, C.; MEMBER, S. ; TAO, D. **Robust Face Recognition via Multimodal Deep**. IEEE Transactions on Multimedia, 17(11):1–10, 2015.
- FACELI, K. **Inteligência Artificial**. LTC, 2011.
- FADHIL AHMED, D.; KHALAF, A. H. **Artificial Neural Networks Controller for Crude Oil Distillation Column of Baiji Refinery**. Journal of Chemical Engineering & Process Technology, 07(01):1–9, 2015.
- FU, Y.; GUO, G. ; HUANG, T. S. **Age synthesis and estimation via faces: A survey**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(11):1955–1976, 2010.
- FUKUSHIMA, K. **Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position**. Biological Cybernetics, 36(4):193–202, 1980.
- GHAZNAVI, A.; ALIYARI, M. ; MOHAMMADI, M. R. **Predicting Stock Price Changes of Tehran Artmis Company Using Radial Basis Function Neural Networks**. 10(8):972–978, 2016.
- GONFAUS, J. M.; PEDERSOLI, M.; GONZÁLEZ, J.; VEDALDI, A. ; ROCA, F. X. **Factorized appearances for object detection**. Computer Vision and Image Understanding, 138:92–101, 2015.
- H. FARID. **A Survey of Image Forgery Detection**. IEEE Signal Process. Mag., 26(2):16–25, 2009.
- HAN, K.; YU, D. ; TASHEV, I. **Speech Emotion Recognition Using Deep Neural Network and Extreme Learning Machine**. Fifteenth Annual Conference of . . . , (September):223–227, 2014.
- HEBB, D. O. **The Organization of Behavior: A Neuropsychological Theory**, volumen 63. JOHN WHAT & SONS. INC, 1950.
- HINTON, G. E.; SRIVASTAVA, N. ; SWERSKY, K. **Lecture 6a- overview of mini-batch gradient descent**. COURSERA: Neural Networks for Machine Learning, p. 31, 2012.

- HJELMÅS, E.; LOW, B. K. **Face Detection: A Survey**. Computer Vision and Image Understanding, 83(3):236–274, 2001.
- HOPFIELD, J. J. **Neural networks and physical systems with emergent collective computational abilities**. Proceedings of the National Academy of Sciences, 79(8):2554–2558, 1982.
- HU, G.; PENG, X.; YANG, Y.; HOSPEDALES, T. ; VERBEEK, J. **Frankenstein: Learning Deep Face Representations using Small Data**. IEEE Transactions on Image Processing, 27(1):293–303, 2016.
- JANG, D.-W.; PARK, R.-H. ; MEMBER, S. **Color Fringe Correction by the Color Difference Prediction Using the Logistic Function**. IEEE Transactions on Image Processing, 26(5):2561–2570, 2017.
- JONES, M.; VIOLA, P. **Fast Multi-view Face Detection**. Mitsubishi Electric Research Lab TR2000396, (July), 2003.
- KALAL, Z.; MIKOLAJCZYK, K. ; MATAS, J. **Face-TLD: Tracking-learning-detection applied to faces**. Proceedings - International Conference on Image Processing, ICIP, (i):3789–3792, 2010.
- KEMELMACHER-SHLIZERMAN, I.; SHECHTMAN, E.; GARG, R. ; SEITZ, S. M. **Exploring photobios**. ACM SIGGRAPH 2011 papers on - SIGGRAPH '11, p. 1, 2011.
- KIM, T.-K.; CIPOLLA, R. **MCBoost: Multiple Classifier Boosting for Perceptual Co-clustering of Images and Visual Features**. Advances in Neural Information Processing Systems 21, (January 2008):841–848, 2009.
- KOHONEN, T. **Self-Organized Formation of Topologically Correct Feature Maps**. Biological Cybernetics, 79(8):43–59, 1982.
- KUME, H.; SATO, T. ; YOKOYA, N. **Bundle adjustment using aerial images with two-stage geometric verification**. Computer Vision and Image Understanding, 138:74–84, 2015.
- LAURENTINI, A.; BOTTINO, A. **Computer analysis of face beauty: A survey**. Computer Vision and Image Understanding, 125:184–199, 2014.
- LAWRENCE, S.; TSOI, C. T. ; BACK, A. D. **Function Approximation with Neural Networks and Local Methods: Bias, Variance and Smoothness**. Pattern Recognition, p. 16–21, 1996.
- LECUN, Y.; BOSER, B.; DENKER, J. S.; HENDERSON, D.; HOWARD, R. E.; HUBBARD, W. ; JACKEL, L. D. **Backpropagation Applied to Handwritten Zip Code Recognition**, 1989.

- LECUN, Y.; BENGIO, Y. **Convolutional Networks for Images, Speech, and Time-Series**. The Handbook of Brain Theory and Neural Networks, (May 2013):255–257, 1995.
- LI, S. Z.; ZHU, L.; ZHANG, Z.; BLAKE, A.; ZHANG, H. ; SHUM, H. **Statistical learning of multi-view face detection**. Proc. European Conf. on Computer Vision, p. 67–81, 2002.
- LI, Z.; ZHENG, J. **Single Image De-Hazing Using Globally**. IEEE Transactions on Image Processing, 27(1):442–450, 2018.
- LIN, S.-H. **An Intro[1] S. Lin, “An Introduction Of Computer Security**. Informing Science, (1995):1–7, 2000.
- LIN, Y.-Y. L. Y.-Y.; LIU, T.-L. L. T.-L. **Robust face detection with multi-class boosting**. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), 1, 2005.
- LUKAC, R.; PLATANIOTIS, K. N. **Universal demosaicking for imaging pipelines with an RGB color filter array**. Pattern Recognition, 38(11):2208–2212, 2005.
- LUKÁŠ, J.; FRIDRICH, J. ; GOLJAN, M. **Detecting digital image forgeries using sensor pattern noise**. PROCEEDINGS OF SPIE, THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING, p. 60720Y, 2006.
- MALSBURG, C. **Self-organization of orientation sensitive cells in the striate cortex**. Kybernetik, 14(2):85–100, 1973.
- MCCULLOCH, W. S.; PITTS, W. **A logical calculus of the ideas immanent in nervous activity**. The Bulletin of Mathematical Biophysics, 5(4):115–133, 1943.
- MEVA, D. T.; PHD, C. K. K. ; KOTHARI, A. D. **The Study of Adoption of Neural Network Approach in Fingerprint Recognition**. International Journal of Computer Applications, 40(11):975–8887, 2012.
- MICHAEL, L. **Simultaneous Learning and Prediction**. Turing, p. 348–357, 2012.
- MILLER, O.; PIKAZ, A. ; AVERBUCH, A. **Objects based change detection in a pair of gray-level images**. Pattern Recognition, 38(11):1976–1992, 2005.
- MINSKY, M. L.; PAPERT, S. A. **Perceptron**, 1969.
- MOODY, J. **Fast Learning in Multi-Resolution Hierarchies**. Advances in Neural Information Processing Systems 1, p. 29–39, 1989.

- MOSLEH, A.; SOLA, Y. E.; ZARGARI, F.; MEMBER, S.; ONZON, E. ; LANGLOIS, J. M. P. **Explicit Ringing Removal in Image Deblurring**. IEEE Transactions on Image Processing, 27(2):580–593, 2018.
- NAYAR, S. K. **Attribute and Simile Classifiers for Face Verification**. Compute, (Iccv), 2009.
- NIE, Y.; SU, T.; ZHANG, Z.; SUN, H. ; LI, G. **Dynamic Video Stitching via Shakiness Removing**. IEEE Transactions on Image Processing, 27(1):164–178, 2017.
- NIU, Z.; HUA, G.; WANG, L. ; GAO, X. **Knowledge Based Topic Model for Unsupervised Object Discovery and Localization**. IEEE Transactions on Image Processing, 27(1):50–63, 2017.
- PANTIE, M.; ROTHKRANTZ, L. **Automatic analysis of facial expressions: the state of the art**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12):1424–1445, 2000.
- PENG, C.; GAO, X.; MEMBER, S.; WANG, N. ; LI, J. **Supapixel-Based Face Sketch Photo Synthesis**. IEEE Transactions on Circuits and Systems for Video Technology(TCSVT), 27(2):288–299, 2017.
- QIAN, P.; JIANG, Y.; WANG, S.; SU, K. H.; WANG, J.; HU, L. ; MUZIC, R. F. **Affinity and Penalty Jointly Constrained Spectral Clustering with All-Compatibility, Flexibility, and Robustness**. IEEE Transactions on Neural Networks and Learning Systems, 28(5):1123–1138, 2017.
- QIN, S. J. **Neural Networks for Intelligent Sensors and Control - Practical Issues and Some Solutions**. Neural Systems for Control, (i):215–236, 1997.
- RAMA VARIOR, R.; WANG, G.; LU, J. ; LIU, T. **Learning Invariant Color Features for Person Reidentification**. IEEE Transactions on Image Processing, 25(7):3395–3410, 2016.
- ROCHA, R. H. S. **Reconhecimento de Objetos por Redes Neurais Convolutivas**. p. 49, 2015.
- ROSENBLATT, F. **The Perceptron - A Perceiving and Recognizing Automaton**, 1957.
- RUMELHART, D. E.; HINTON, G. E. ; WILLIAMS, R. J. **Learning internal representations by error propagation**. COGNITIVE SCIENCE, 2(V):34, 1985.

- RUMELHART, D. E.; HINTON, G. E. ; WILLIAMS, R. J. **Learning representations by back-propagating errors**. *Nature*, 323(6088):533–536, 1986.
- SAMAL, A.; IYENGAR, P. A. **Automatic recognition and analysis of human faces and facial expressions: a survey**. *Pattern Recognition*, 25(1):65–77, 1992.
- SCHAUL, T.; ZHANG, S. ; LECUN, Y. **No More Pesky Learning Rates**. 2012.
- SCHROFF, F.; PHILBIN, J. **FaceNet: A Unified Embedding for Face Recognition and Clustering**. In: *CVPR - IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, volumen 1, p. 815–823, 2015.
- SEEMANN, E.; LEIBE, B. ; SCHIELE, B. **Multi-aspect detection of articulated objects**. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:1582–1588, 2006.
- SHAN, Y.; HAN, F.; SAWHNEY, H. S. ; KUMAR, R. **Learning exemplar-based categorization for the detection of multi-view multi-pose objects**. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:1431–1438, 2006.
- SUN, Y.; WANG, X. ; TANG, X. **Deeply Learned Face Representations Are Sparse, Selective, and Robust**. In: *CVPR - IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, p. 2892–2900, 2015.
- SUTTON, R. S. **Temporal credit assignment in reinforcement learning**. PhD thesis, University of Massachusetts Amherst, 1984.
- SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCKE, V.; RABINOVICH, A.; HILL, C. ; ARBOR, A. **Going Deeper with Convolutions**. In: *CVPR - IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, p. 1–9, 2014.
- TABERNIK, D.; LEONARDIS, A.; BOBEN, M.; SKOČAJ, D. ; KRISTAN, M. **Adding discriminative power to a generative hierarchical compositional model using histograms of compositions**. *Computer Vision and Image Understanding*, 138:102–113, 2015.
- TAIGMAN, Y.; YANG, M.; RANZATO, M. ; WOLF, L. **DeepFace: Closing the gap to human-level performance in face verification**. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 1701–1708, 2014.

- TIRUNAGARI, S.; POH, N.; WINDRIDGE, D.; IORLIAM, A.; SUKI, N. ; HO, A. T. **Detection of face spoofing using visual dynamics**. IEEE Transactions on Information Forensics and Security, 10(4):762–777, 2015.
- TSENG, C. C.; LIEN, J. J. J. **Colored exaggerative caricature creation using inter- and intra-correlations of feature shapes and positions**. Image and Vision Computing, 30(1):15–25, 2012.
- TU, Z. **Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering**. Proceedings of the IEEE International Conference on Computer Vision, II:1589–1596, 2005.
- TUNG, F.; LITTLE, J. J. **Improving scene attribute recognition using web-scale object detectors**. Computer Vision and Image Understanding, 138:86–91, 2015.
- ULLRICH, K. **Boundary Detection in Music Structure Analysis Using Convolutional Neural Networks**. Proc. of the 14th International Society for Music Information Retrieval Conference, 2014.
- VATSARAJ, M. S. **Anomalous Behaviour Detection in Crowded Environments Using Classifiers Artificial Neural Network and Support Vector Machine**. 3:654–661, 2017.
- VIOLA, P.; JONES, M. **Rapid object detection using a boosted cascade of simple features**. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 1:I–511–I–518, 2001.
- WANG, J.; SONG, Y.; LEUNG, T.; ROSENBERG, C.; WANG, J.; PHILBIN, J.; CHEN, B. ; WU, Y. **Learning fine-grained image similarity with deep ranking**. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 1386–1393, 2014.
- WANG, M.; LI, W.; LIU, D.; NI, B.; SHEN, J. ; YAN, S. **Facilitating Image Search with a Scalable and Compact Semantic Mapping**. IEEE Transactions on Cybernetics, 45(8):1561–1574, 2015.
- WANG, N.; TAO, D.; GAO, X.; LI, X. ; LI, J. **Transductive face sketch-photo synthesis**. IEEE Transactions on Neural Networks and Learning Systems, 24(9):1364–1376, 2013.
- WANG, Y.; LIU, Z.; HUA, G.; WEN, Z.; ZHANG, Z. ; SAMARAS, D. **Face re-lighting from a single image under harsh lighting conditions**. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 31(11):1968–1984, 2007.

- WIENER, N. **Cybernetics or Control and Communication in the Animal and the Machine.** In: M.I.T. PRESS, p. 109–124. THE M.I.T. PRESS Cambridge, Massachusetts, second edi edition, 1948.
- WU, B.; NEVATIA, R. **Cluster boosted tree classifier for multi-view, multi-pose object detection.** Proceedings of the IEEE International Conference on Computer Vision, 2007.
- WU, S.; HONG, L. **Hand tracking in a natural conversational environment by the interacting multiple model and probabilistic data association (IMM-PDA) algorithm.** Pattern Recognition, 38(11):2143–2158, 2005.
- YUAN, Y.; FANG, J. ; WANG, Q. **Online anomaly detection in crowd scenes via structure analysis.** IEEE Transactions on Cybernetics, 45(3):562–575, 2015.
- ZAFEIRIOU, S.; ZHANG, C. ; ZHANG, Z. **A survey on face detection in the wild: Past, present and future.** Computer Vision and Image Understanding, 138:1–24, 2015.
- ZHANG, C.; ZHANG, Z. **Winner-take-all multiple category boosting for multi-view face detection.** Communication, 2009.
- ZHANG, S.; YANG, H. ; YIN, Z. P. **Transferred deep convolutional neural network features for extensive facial landmark localization.** IEEE Signal Processing Letters, 23(4):478–482, 2016.
- ZHANG, Y.; DONG, W.; MA, C.; MEI, X.; LI, K.; HUANG, F.; HU, B. G. ; DEUSSEN, O. **Data-Driven Synthesis of Cartoon Faces Using Different Styles.** IEEE Transactions on Image Processing, 26(1):464–478, 2017.
- ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P. ; ROSENFELD, A. **Face recognition: A literature survey.** Acm Computing Surveys, 35(4):399–458, 2003.
- ZHU, Q.; SHAO, L.; LI, X. ; WANG, L. **Targeting Accurate Object Extraction From an Image: A Comprehensive Study of Natural Image Matting.** IEEE Transactions on Neural Networks and Learning Systems, 26(2):185–207, 2015.
- ZHU, Z.; LUO, P.; WANG, X. ; TANG, X. **Recover Canonical-View Faces in the Wild with Deep Neural Networks.** CoRR, p. 1–10, 2014.