

OCR for document processing

Computer Vision Final Defense



Table of contents

01 Introduction

Approach to solve the problem of document processing

02 Program pipeline

3 main steps: Pre-processing, Text Detection, and Text Recognition

03 Theoretical foundation

Theoretical background of the techniques used in our system

04 Result and Conclusion

The conclusion of the project



01

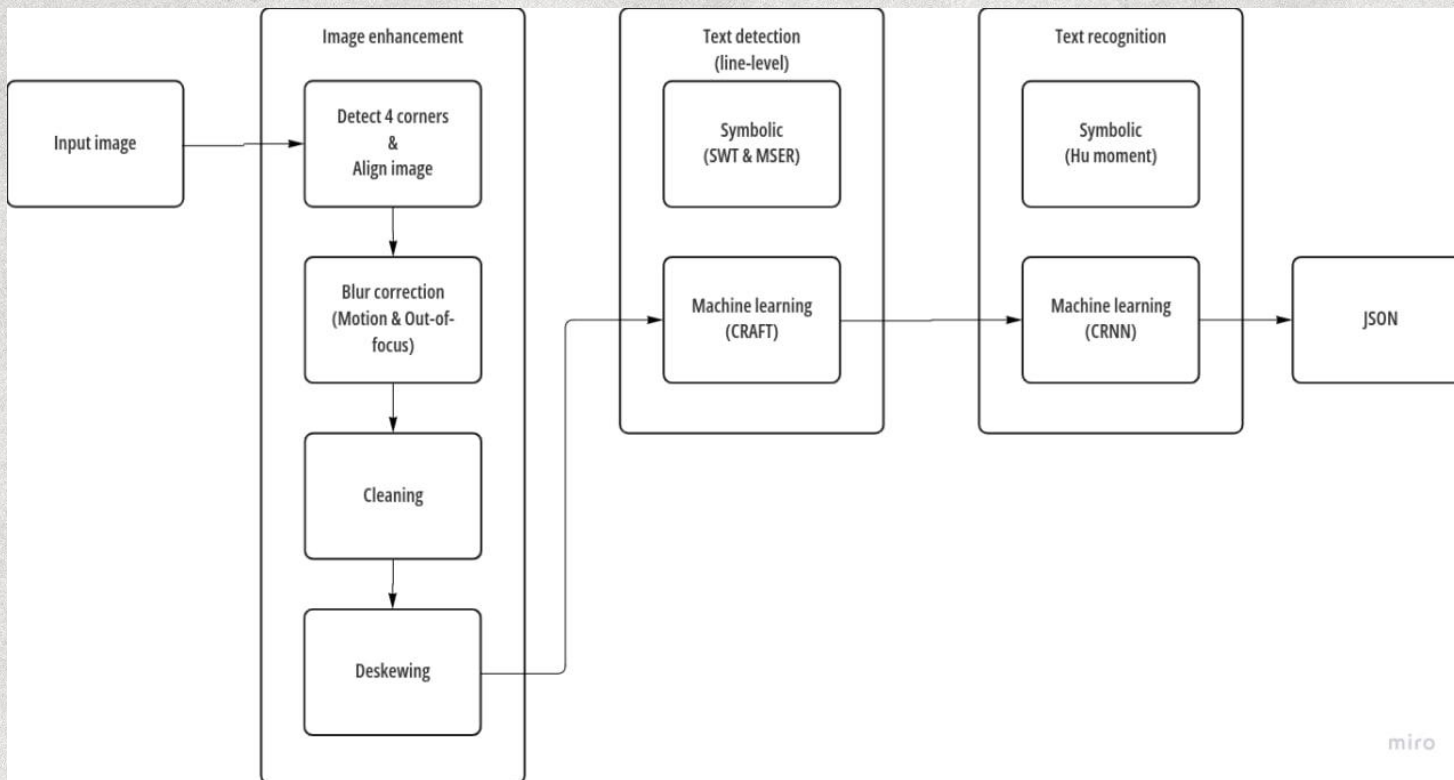
Introduction

A simple OCR solution for
document processing





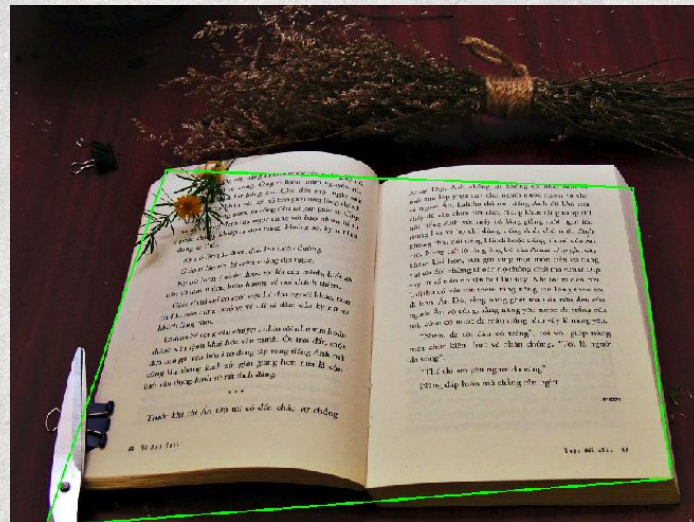
Our design system pipeline



Corner detection and image alignment

Our methods as follows:

- Median blur filter
- K-means algorithm
- Connected component finding algorithm, approximate it by a convex quadrilateral





Blur correction

Fourier transform and Wiener filter

GIẤY CHỨNG NHẬN ĐĂNG KÝ DOANH NGHIỆP
CÔNG TY CỔ PHẦN

Mã số doanh nghiệp: 0309829579

Đăng ký lần đầu: ngày 11 tháng 03 năm 2010

Đăng ký thay đổi lần thứ: 7, ngày 29 tháng 11 năm 2017

1. Tên công ty

Tên công ty viết bằng tiếng Việt: CÔNG TY CỔ PHẦN TRUNG TÂM CHĂM SÓC
SỨC KHỎE CỘNG ĐỒNG CHẠCTên công ty viết bằng tiếng nước ngoài: CHAC COMMUNITY HEALTHCARE
CENTER CORPORATION

Tên công ty viết tắt: TRUNG TÂM CHĂM SÓC SỨC KHỎE CỘNG ĐỒNG CHẠC

2. Địa chỉ trụ sở chính

110A Ngô Quyền, Phường 08, Quận 5, Thành phố Hồ Chí Minh, Việt Nam

Điện thoại: 0902714743

Fax:

Email:

Website:

3. Vốn điều lệ

Vốn điều lệ: 8.000.000.000 đồng

Bảng chữ: Tám tỷ đồng

Mệnh giá cổ phần: 10.000 đồng

Tổng số cổ phần:

4. Người đại diện theo pháp luật của công ty

* Họ và tên: NGÔ HỒN ĐÌNH

Giới tính: Nam

Chức danh: Giám đốc

Sinh ngày: 25/07/1984

Dân tộc: Hoa

Quốc tịch: Việt Nam

Loại giấy tờ chứng thực cá nhân: Chứng minh nhân dân

Số giấy chứng thực cá nhân: 023692354

Ngày cấp: 23/06/2005

Nơi cấp: Công an Thành phố Hồ Chí Minh

Nơi đăng ký hộ khẩu thường trú: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố
Hồ Chí Minh, Việt NamChỗ ở hiện tại: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố Hồ Chí Minh,
Việt NamTRƯỞNG PHÒNG
Phó Trưởng Phòng

Hoàng Anh Phan

GIẤY CHỨNG NHẬN ĐĂNG KÝ DOANH NGHIỆP
CÔNG TY CỔ PHẦN

Mã số doanh nghiệp: 0309829579

Đăng ký lần đầu: ngày 11 tháng 03 năm 2010

Đăng ký thay đổi lần thứ: 7, ngày 29 tháng 11 năm 2017

1. Tên công ty

Tên công ty viết bằng tiếng Việt: CÔNG TY CỔ PHẦN TRUNG TÂM CHĂM SÓC
SỨC KHỎE CỘNG ĐỒNG CHẠCTên công ty viết bằng tiếng nước ngoài: CHAC COMMUNITY HEALTHCARE
CENTER CORPORATION

Tên công ty viết tắt: TRUNG TÂM CHĂM SÓC SỨC KHỎE CỘNG ĐỒNG CHẠC

2. Địa chỉ trụ sở chính

110A Ngô Quyền, Phường 08, Quận 5, Thành phố Hồ Chí Minh, Việt Nam

Điện thoại: 0902714743

Fax:

Email:

Website:

3. Vốn điều lệ

Vốn điều lệ: 8.000.000.000 đồng

Bảng chữ: Tám tỷ đồng

Mệnh giá cổ phần: 10.000 đồng

Tổng số cổ phần:

4. Người đại diện theo pháp luật của công ty

* Họ và tên: NGÔ HỒN ĐÌNH

Giới tính: Nam

Chức danh: Giám đốc

Sinh ngày: 25/07/1984

Dân tộc: Hoa

Quốc tịch: Việt Nam

Loại giấy tờ chứng thực cá nhân: Chứng minh nhân dân

Số giấy chứng thực cá nhân: 023692354

Ngày cấp: 23/06/2005

Nơi cấp: Công an Thành phố Hồ Chí Minh

Nơi đăng ký hộ khẩu thường trú: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố
Hồ Chí Minh, Việt NamChỗ ở hiện tại: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố Hồ Chí Minh,
Việt NamTRƯỞNG PHÒNG
Phó Trưởng Phòng

Hoàng Anh Phan

Cleaning

20
22

Tên công ty viết bằng tiếng Việt: CÔNG TY CỔ PHẦN TRUNG TÂM CHĂM SÓC SỨC KHỎE CỘNG ĐỒNG
SỨC KHỎE CỘNG ĐỒNG CHAC
Tên công ty viết bằng tiếng nước ngoài: CHAC COMMUNITY HEALTH CARE
CENTER CORPORATION
Tên công ty viết tắt: TRUNG TÂM CHĂM SÓC SỨC KHỎE CỘNG ĐỒNG
2. Địa chỉ trụ sở chính
110A Ngô Quyền, Phường 08, Quận 5, Thành phố Hồ Chí Minh, Việt Nam
Điện thoại: 0902714743
Email:
3. Vốn điều lệ
Vốn điều lệ: 8.000.000.000 đồng
Bằng chữ: Tám tỷ đồng
Mệnh giá cổ phần: 10.000 đồng
Tổng số cổ phần:
4. Người đại diện theo pháp luật của công ty
* Họ và tên: NGÔ HÓN ĐÌNH
Chức danh: Giám đốc
Sinh ngày: 25/07/1984
Loại giấy tờ chứng thực cá nhân: Chứng minh nhân dân
Số giấy chứng thực cá nhân: 023692354
Ngày cấp: 23/06/2005
Nơi đăng ký hộ khẩu thường trú: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố Hồ Chí Minh, Việt Nam
Chỗ ở hiện tại: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố Hồ Chí Minh, Việt Nam
Giới tính: Nam
Quốc tịch: Việt Nam

Deskewing

1. Tên công ty
Tên công ty viết bằng tiếng Việt: CÔNG TY CỔ PHẦN TRUNG TÂM CHĂM SÓC SỨC KHỎE CỘNG ĐỒNG CHAC
Tên công ty viết bằng tiếng nước ngoài: CHAC COMMUNITY HEALTH CARE
CENTER CORPORATION
Tên công ty viết tắt: TRUNG TÂM CHĂM SÓC SỨC KHỎE CỘNG ĐỒNG
2. Địa chỉ trụ sở chính
110A Ngô Quyền, Phường 08, Quận 5, Thành phố Hồ Chí Minh, Việt Nam
Điện thoại: 0902714743
Email:
3. Vốn điều lệ
Vốn điều lệ: 8.000.000.000 đồng
Bằng chữ: Tám tỷ đồng
Mệnh giá cổ phần: 10.000 đồng
Tổng số cổ phần:
4. Người đại diện theo pháp luật của công ty
* Họ và tên: NGÔ HÓN ĐÌNH
Chức danh: Giám đốc
Sinh ngày: 25/07/1984 Dân tộc: Hoa
Loại giấy tờ chứng thực cá nhân: Chứng minh nhân dân
Số giấy chứng thực cá nhân: 023692354
Ngày cấp: 23/06/2005 Nơi cấp: Công an Thành phố Hồ Chí Minh
Nơi đăng ký hộ khẩu thường trú: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố Hồ Chí Minh, Việt Nam
Chỗ ở hiện tại: 1442/46 đường 3/2, Phường 2, Quận 11, Thành phố Hồ Chí Minh, Việt Nam

20
22

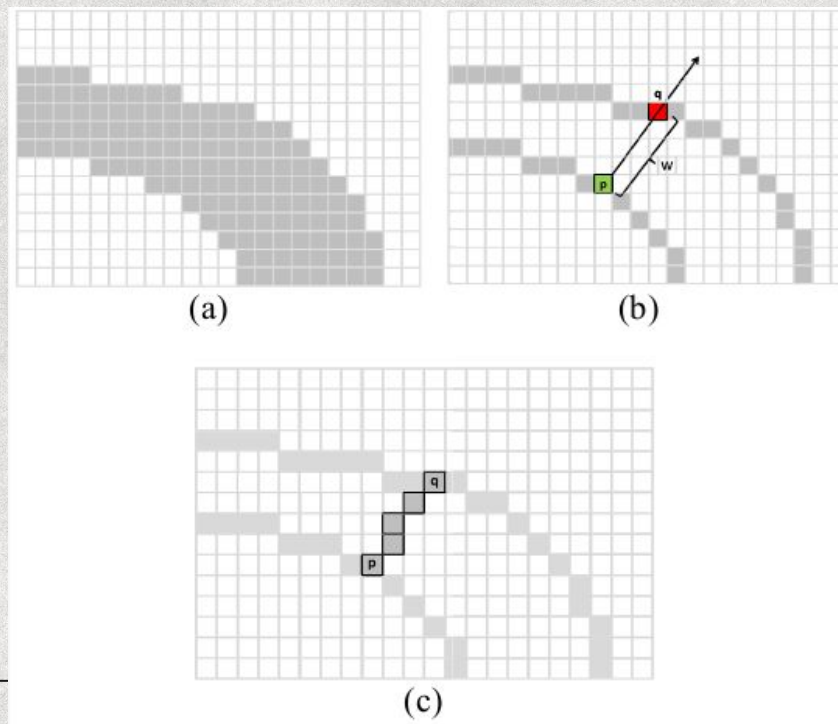


Text detection

Identify the areas that contain text in the image:
Character-level, word-level, or line-level



Symbolic method: SWT

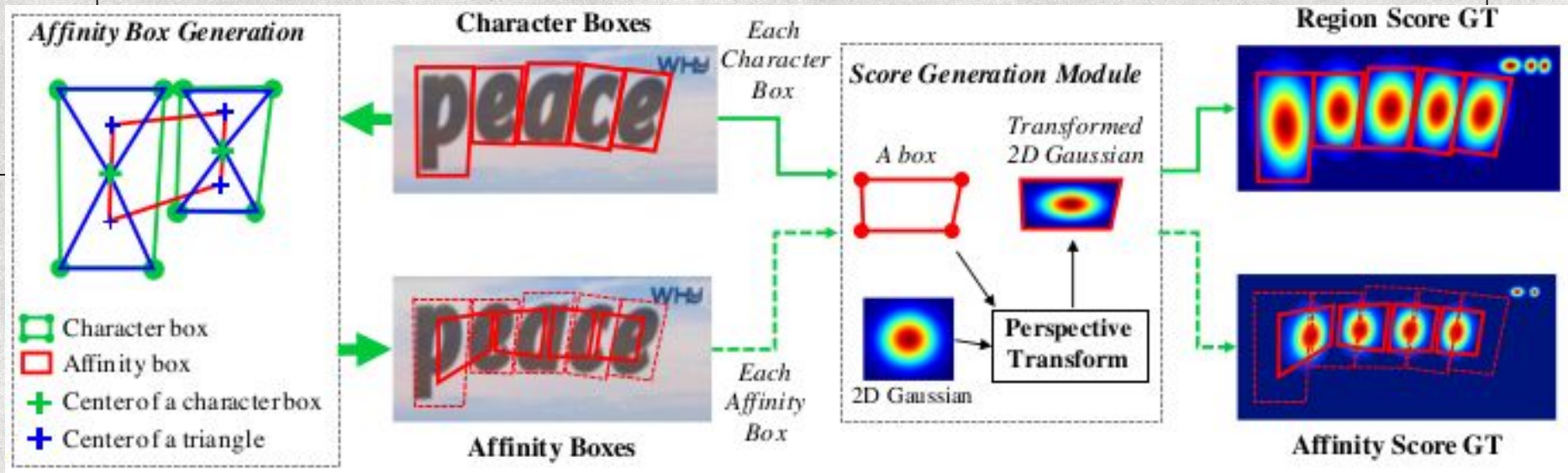


Deep learning method

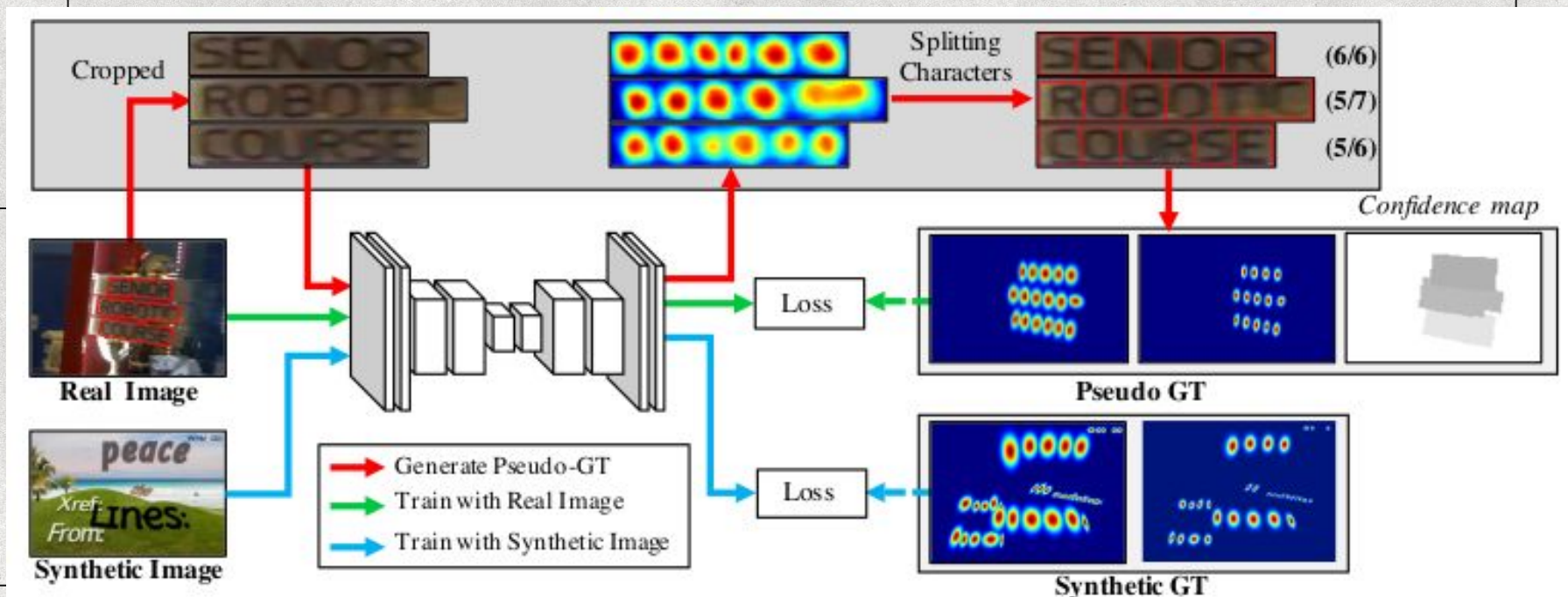
CRAFT, CTPN, EAST, FOTS, PixelLink, PSENet.

CTPN, EAST, FOTS, PixelLink: text areas that **overlap** each other.
CTPN, EAST, FOTS, PixelLink, PSENet: falsely detect **non-text** areas

CRAFT



CRAFT





Text recognition

Take the text area of the previous stage as input and outputs a JSON file containing the coordinates and contents of the corresponding text.



Symbolic method

ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789

Shape matching



The predicted text is TdA8lOd2TlOlOUd





Deep learning method

CNN backbone

VGG: ~22 million params

MobileNetV3: ~7 million params

The second model not run yet, so we temporarily use pretrained

	Transformer	Seq2seq
Total params	17,363,968	2,204,393
Inference time	0.163s	0.03s



Theoretical foundation

Theoretical background of the techniques used in our system. That is the knowledge that has not been taught in the IT4343E - Computer Vision class.



Convolutional theorem

Spatial domain	Frequency domain
$g(x) = f(x) * h(x)$ convolution	$G(u) = F(u)H(u)$ multiplication
$g(x) = f(x)h(x)$ multiplication	$G(u) = F(u) * H(u)$ convolution

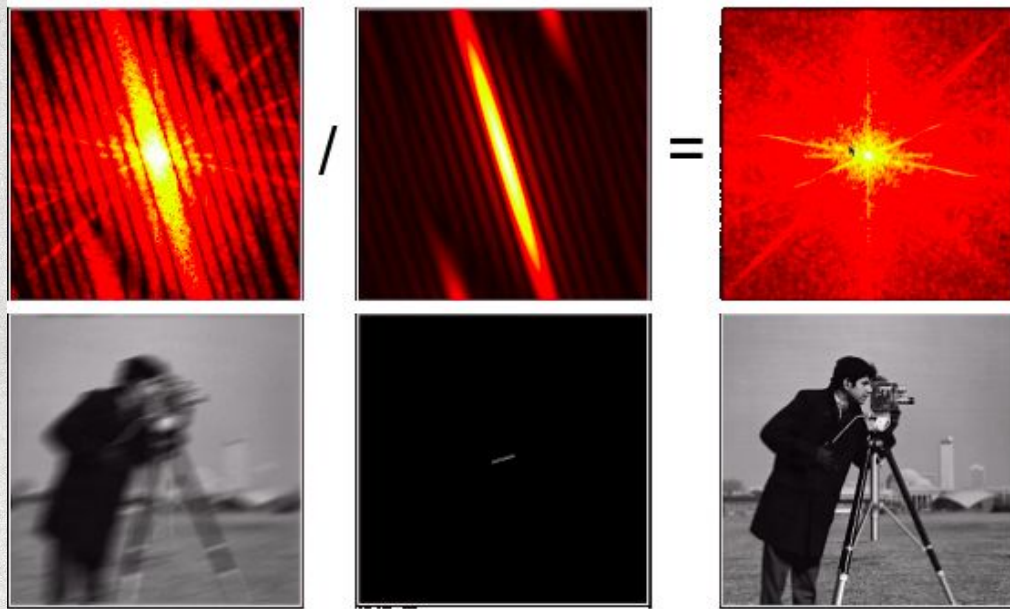
Motion blur



Figure 3.1: Camera shake and deconvolution. From left to right: $f(x,y)$, $h(x,y)$, $g(x,y)$ where $f(x,y) * h(x,y) = g(x,y)$.

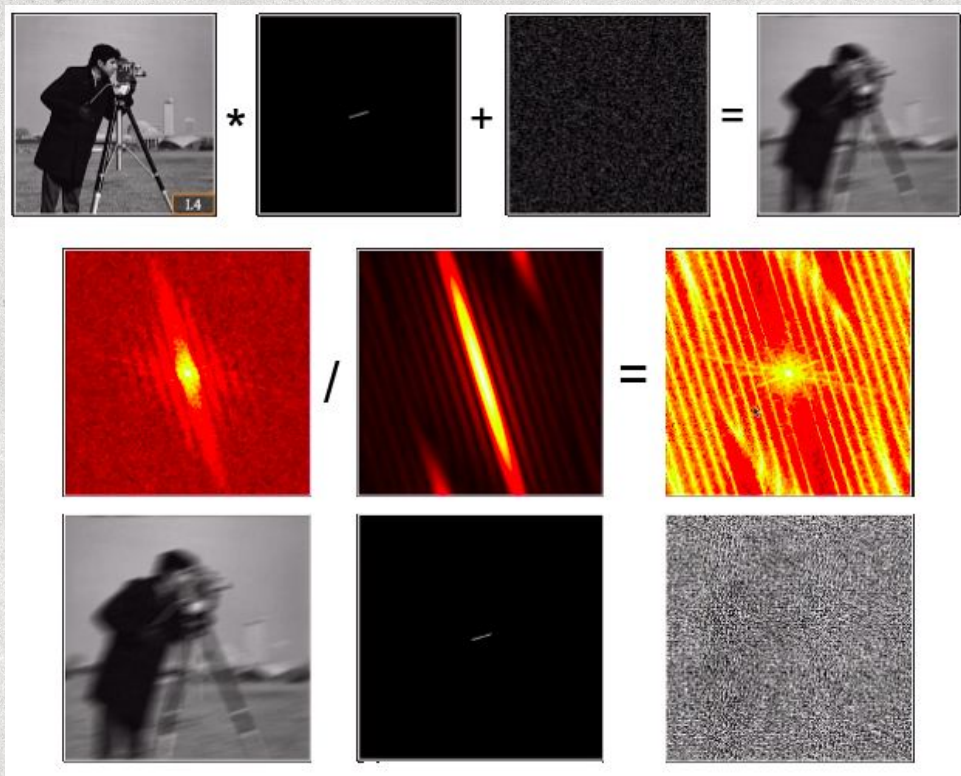


Deconvolution



Deconvolution for blur correction. From left to right: $f(x, y)$, $h(x, y)$, $g(x, y)$ where $f(x, y) * h(x, y) = g(x, y)$

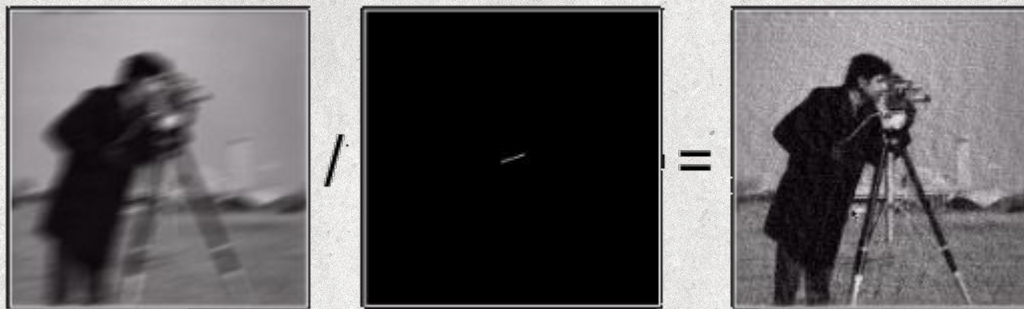
Naive deconvolution with noise

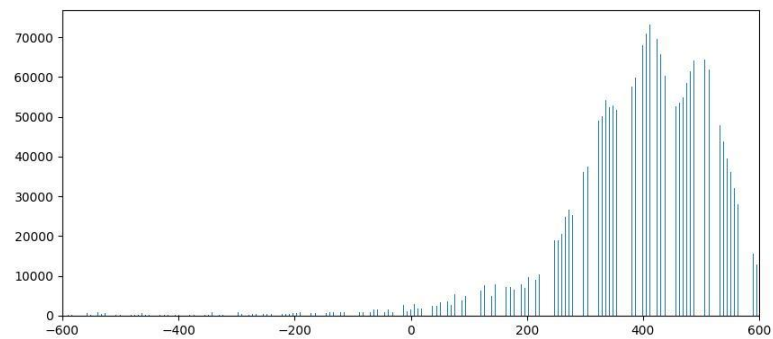
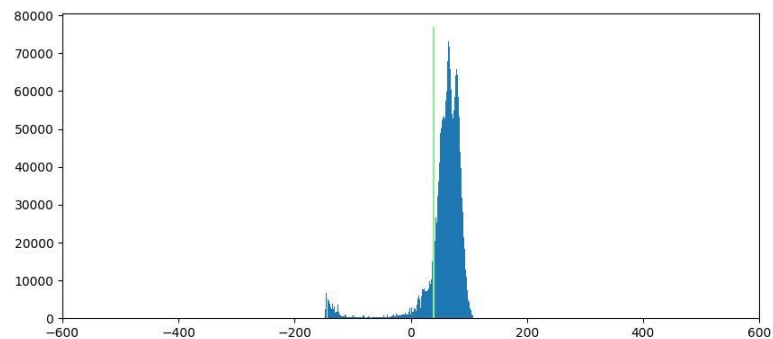
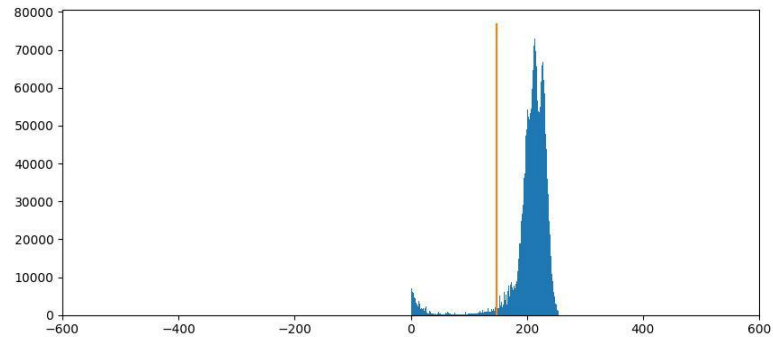




Wiener filter

$$F(u, v) = \frac{G(u, v)}{H(u, v)} \left[\frac{1}{1 + \frac{NSR(u, v)}{|H(u, v)|^2}} \right]$$





Clean

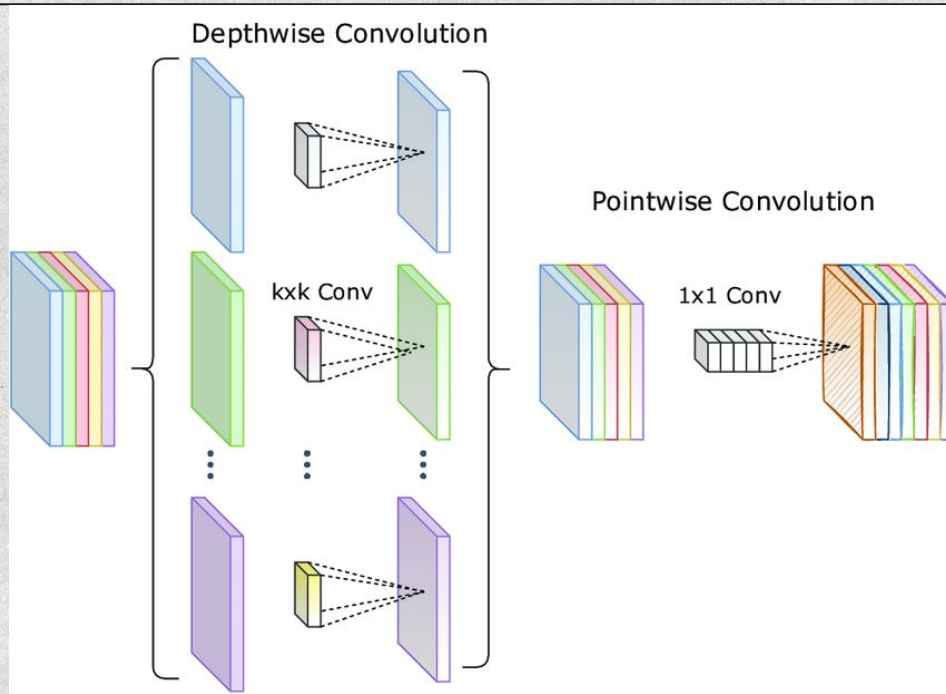




MobileNet V3

For CRNN backbone



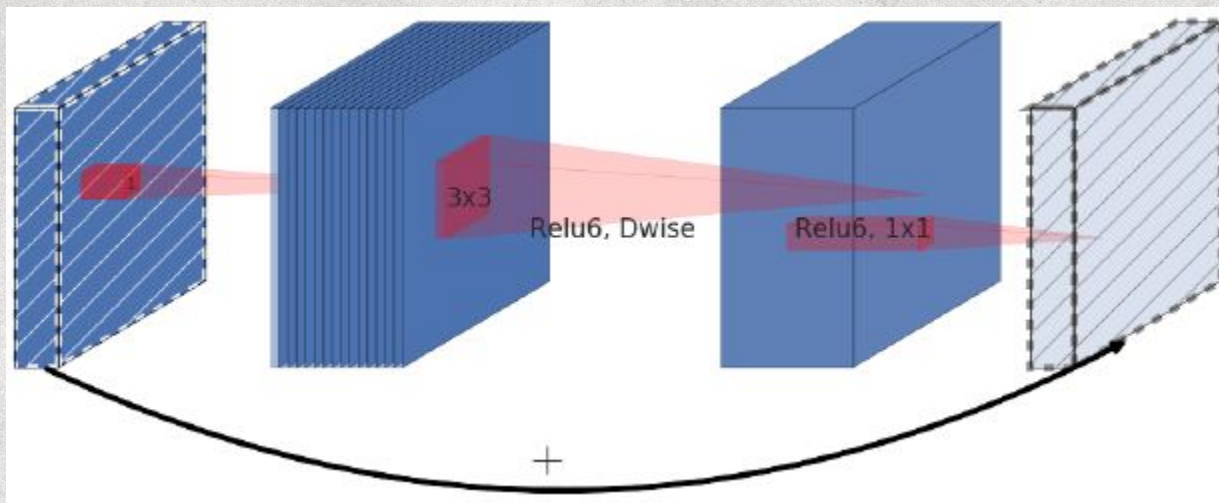


Depthwise separable convolutions

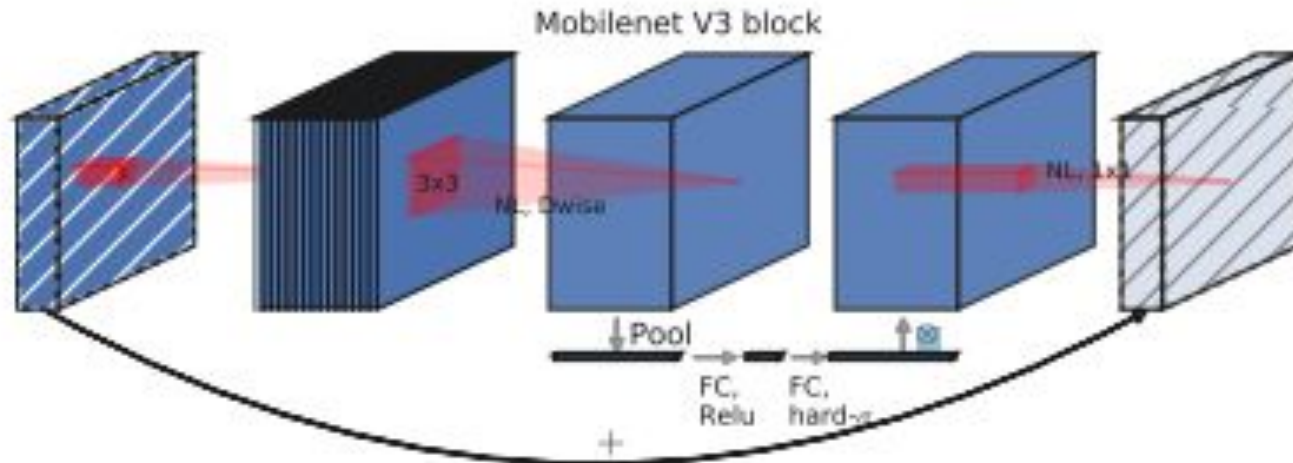
In general, we have the formular $\frac{\text{cost}_{\text{depwise separable}}}{\text{cost}_{\text{normal}}} = \frac{1}{n_{c'}} + \frac{1}{f^2}$ where c' is the number of output channels and f is the kernel size.



MobileNetV2 building block



MobileNet V3 building block





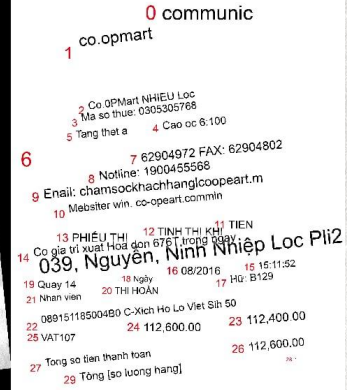
04

Result and Conclusion

We have made a relatively complete product. Although, it still hasn't lived up to our original expectations



Some results



GIẤY KHAI SINH

(BẢN CHÍNH)

Họ và tên: NGUYỄN NAM PHƯƠNG Giới tính: Nữ
 Ngày, tháng, năm sinh: 08/16/2011 Ghi bằng chữ: ngày mùng tám tháng sáu - năm hai nghìn không trăm mười một
 Nơi sinh: Bệnh viện Việt - Pháp Hà Nội
 Dân tộc: Kinh Quốc tịch: Việt Nam
 Họ và tên cha: Nguyễn Ninh Hồng Quang
 Dân tộc: Kinh Quốc tịch: Việt Nam Năm sinh: 1980
 Nơi thường trú/tạm trú: 309 nhà E2 Bạch Thoa - Hai Bà Trưng - Hà Nội
 Họ và tên mẹ: Phạm Thủy Trang
 Dân tộc: Kinh Quốc tịch: Việt Nam Năm sinh: 1989
 Nơi thường trú/tạm trú: 309 nhà E2 Bạch Thoa - Hai Bà Trưng - Hà Nội
 Nơi đăng ký: UBND phường Bạch Thoa - Quận Hai Bà Trưng - Hà Nội
 Ngày, tháng, năm đăng ký: 05/18/2011
 Ghi chú:
 Họ và tên người đi khai sinh: Nguyễn Ninh Hồng Quang
 Quan hệ với người được khai sinh: cha

NGƯỜI THỰC HIỆN
(Ký, ghi rõ họ tên)

Nguyễn Thị Thu Hiền

NGƯỜI KÝ GIẤY KHAI SINH
(Ký, ghi rõ họ tên, chức vụ và đóng dấu)



CHỦ TỊCH

Nguyễn Thị Kim Hoa

GIẤY KHAI SINH

GIẤY KHAI SINH

(BẢN CHÍNH)

(BẢN CHÍNH)

Họ và tên: Nguyễn NAM Phương Giới tính: Nữ
 Ngày, tháng, năm sinh: 08/16/2011 Ghi bằng chữ: ngày mùng tám tháng sáu - năm hai nghìn không trăm mười một
 Nơi sinh: Bệnh viện Việt Pháp Hà Nội
 Dân tộc: Kinh Quốc tịch: Viet Nam
 Họ và tên cha: Nguyễn Ninh Hồng Quang Quốc tịch: Việt Nam
 Dân tộc: Kinh Quốc tịch: Việt Nam Năm sinh: 1980
 Nơi thường trú/tạm trú: 309 nhà E2 Bạch Thoa - Hai Bà Trưng - Hà Nội
 Họ và tên mẹ: Phạm Thủy Trang Quốc tịch: Viet Nam Năm sinh: 1989
 Nơi thường trú/tạm trú: 309 nhà E2 Bạch Thoa - Hai Bà Trưng - Hà Nội
 Nơi đăng ký: UBND phường Bạch Thoa - Quận Hai Bà Trưng - Hà Nội
 Ngày, tháng, năm đăng ký: 05/18/2011
 Ghi chú:
 Họ và tên người đi khai sinh: Nguyễn Ninh Hồng Quang
 Quan hệ với người được khai sinh: cha

NGƯỜI THỰC HIỆN

(Ký, ghi rõ họ tên)

Nguyễn Thị Thu Hiền

NGƯỜI KÝ GIẤY KHAI SINH

(Ký, ghi rõ họ tên, chức vụ và đóng dấu)

Nguyễn Thị Kim Hoa



CHỦ TỊCH

Nguyễn Thị Kim Hoa

Đặc điểm nổi bật của sự phân bố dân cư của các tộc người ở nước ta là tình trạng cư trú xen kẽ nhau nên các tộc người không có lãnh thổ rõ rệt. Căn cứ vào ngôn ngữ, các nhà dân tộc học chia các tộc người thành 8 nhóm như sau:

1. Việt Mường (gồm Việt, Mường, Chứt...)
2. Tày – Thái (gồm Tày, Nùng, Thái, Bố Y, Cao Lan, Sán Chỉ, Lào...)
3. H'mông – Dao (gồm H'mông, Dao, Pà Thẻn)
4. Tạng – Miến (gồm Hà Nhì, Lô Lô, Xá ...)
5. Hán (gồm Hoa, Sán Diu...)
6. Môn – Khơme (Khơ mú, Kháng, Xinh Mun, Hơrê, Xơđăng, Bana, K'ho, Mạ, Rơmăm, Khơme...)
7. Mã Lai – Đa đảo (gồm Chăm, Giarai, Êđê, Raglai...)
8. Hỗn hợp Nam Á (gồm La chỉ, Laha, Pư pẻo ...)

Đặc điểm nổi bật của sự phân bố dân cư của các tộc người ở nước ta là tình trạng cư trú xen kẽ nhau nên các tộc người không có lãnh thổ rõ rệt, Căn cứ vào ngôn ngữ, các nhà dân tộc học chia các tộc người thành 8 nhóm như sau:

1. Việt Mường (gồm Việt, Mường, Chứt...)
2. Tày ? Thái (gồm Tày, Nùng, Thái, Bố Y, Cao Lan, Sán Chỉ, Lào....)
3. H'mông ? Dao (gồm H'mông, Dao, Pà Thẻn)
4. Tạng - Miến (gồm Hà Nhì, Lô Lô, Xá)
5. Hán (gồm hoa, sán diu...)
6. Môn 11 Khơme (Khomú, Kháng, Xinh Mun, Hơrê, Xơđăng, Bana, Kho, Mạ, Rơmăm, Khơme...)
- T. Mã Lai - Đa đảo (gồm Chăm; Giarai, Êđê, Raglai c.) 15 0311 gnob (1911) 1990)
- Biên 8: Hỗn hợp Nam A (gồm La chỉ, Lăha, Pư pẻo....)



Thanks!

Do you have any questions?

