FACULTY OF ENGINEERING AND ARCHITECTURE

Using GANs to optimize Pose Estimation on Art Collections

Tristan Verheecke
Student number: 20043518

Supervisors: Prof. dr. ir. Dieter De Witte, Prof. dr. Steven Verstockt
Counsellor: Kenzo Milleville

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Information Engineering Technology

Academic year 2022-2023

GHENT UNIVERSITY

Preface

I've been interested in Art my entire life. In fact, I've a degree in the Fine Arts from LUCA School of Arts. There, I was known for my technical ability and one of my professors at the time asked me why I didn't do anything with that in my artworks. That remark has since stuck with me and was part of my motivation to apply for readmission for my Master of Science. With all the advancements in AI, I started thinking more and more about doing work with that. Like Matisse and Turner, I'm not satisfied with the tools available, but want to create my own.

It was therefor to my delight that I was able to work on this thesis which has provided me the opportunity to acquire more insight in the subject. I would like to thank my supervisors Dieter De Witte and Steven Verstockt for this wonderful opportunity, and my counsellor Kenzo Milleville for his great guidance. As well as all the other people at IDLab for their feedback. I also want to thank Karine Lacaracina, Lies Van De Cappelle and the other people at RMFAB for providing help with the artistic sensibilities of the thesis.

Enjoy the read,

Tristan Verheecke
Ghent, June 2023

Conference Paper Title*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

This document is a model and instructions for LaTeX. Please observe the conference page limits.

II. Ease of Use

A. Maintaining the Integrity of the Specifications

The IEEEtran class file is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

Identify applicable funding agency here. If none, delete this.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: “Wb/m^2” or “webers per square meter”, not “webers/m^2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
- Use a zero before decimal points: “0.25”, not “.25”. Use “cm^3”, not “cc”.)

C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \tag{1}$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not

"Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

D. LaTeX-Specific Advice

Please use "soft" (e.g., `\eqref{Eq}`) cross references instead of "hard" references (e.g., `(1)`). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don't use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in LaTeX will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you've discovered a new method of counting.

BibTeX does not work by magic. It doesn't get the bibliographic data from thin air but from .bib files. If you use BibTeX to produce a bibliography you must send the .bib files.

LaTeX can't read your mind. If you assign the same label to a subsubsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

LaTeX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it's supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won't be any anyway) and it might stop a wanted equation number in the surrounding equation.

E. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum μ_0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o".
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset", not an "insert". The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively".
- In your paper title, if the words "that uses" can accurately replace the word "using", capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect", "complement" and "compliment", "discreet" and "discrete", "principal" and "principle".
- Do not confuse "imply" and "infer".
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al.".
- The abbreviation "i.e." means "that is", and the abbreviation "e.g." means "for example".

An excellent style manual for science writers is [7].

F. Authors and Affiliations

The class file is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

G. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is "Heading 5". Use "figure caption" for your Figure captions, and "table head" for your table title. Run-in heads, such as "Abstract", will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

H. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 1", even at the beginning of a sentence.

TABLE I
TABLE TYPE STYLES

Table Head	Table Column Head		
	Table column subhead	*Subhead*	*Subhead*
copy	More table copy[a]		

[a]Sample of a Table footnote.

Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
[4] K. Elissa, "Title of paper if known," unpublished.
[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

Contents

List of Figures

List of Tables

List of Acronyms

A

ASMs	Active Shape Models , 3

C

cGAN	conditional Generative Adversarial Network , 5
CNN	Convolutional Neural Network , 3, 4, 6
CPMs	Convolutional Pose Machines viii, 4, 6, 11
CPN	Cascaded Pyramid Network , 5

G

GAN	Generative Adversarial Network , 5

H

HPE	Human Pose Estimation , 2–6

I

ILP	Integer Linear Programming , 6

M

MPII Max Planck Institute for Informatics viii, 7

N

NMS Non-Maximum-Suppression , 5

P

PAF Part Affinity Field , 6

PAF Part Association Fields , 6

PIF Part Intensity Fields , 6

R

ResNet Residual Network , 5

RMFAB Royal Museums of Fine Arts of Belgium , 1

RPME Regional Multi-person Pose Estimation , 5

S

SAHR Scale-adaptive Heatmap Regression , 6

SMPL Skinned Multi-Person Linear , 3

W

WAHR Weight-adaptive Heatmap Regression , 6

List of Code Fragments

1

Introduction

1.1 Problem definition

To make art collections more accessible, museums put a huge effort in digitalizing their catalogue. However, they don't contain much metadata about the content and it is time-consuming to enhance them manually. To make ths process easier, they want to utilize computer vision. Art collections (paintings, statues, drawings, etc.) turn out to be less interpretable by the algorithms that were developed for photography over the last few decades. These scan the images in search of recognizable objects and add their labels to the metadata. Even the latest state-of-the-art technology, struggles to recognize objects when pointed at a painting in a museum. A solution may be to start over and have paintings annotated by humans.

This has been done in 2 recent projects: Saint-George-On-A-Bike [10] and INSIGHT [11]. However, paintings are very complex and manual annotation doesn't scale and is very expensive. For example, 10,000 paintings were annotated by Royal Museums of Fine Arts of Belgium (RMFAB) with no clear return on investment. They spent a year on this and this is not something they want to repeat. How can we automate this process and ensure that state-of-the-art computer vision models give good results on paintings and artworks?

Specifically for this thesis, pose estimation will be investigated.

2

Literature study

We will first examine the effectiveness of existing models on a collection of paintings from 2 different movements. For this we will need to have a pose estimator, a style transformer and a collection of test data.

A first method: We will first convert the test data with the style transformer to a painting and then we will apply pose estimation. The test data will have coordinates of the joints, which we will compare with the results of the pose estimation. However, the joints are of the original image. How do we convert those coordinates to map to the styled image? Problem: This method does not use any real paintings and will be susceptible to the accuracy of the style transformer.

A second method: We can apply pose estimation to real paintings and then convert them to a realistic image with style transfer. We can then use pose estimation to the realistic images and compare them with the style transformed results. This will also require a way to map the results of the real painting to that of the style transformed. Problem: While we're using real paintings now, the results will still depend on the accuracy of style transformer.

A third method: We can annotate the paintings ourselves and use pose estimation to assess the pose estimation algorithms. Problem: We must annotate the paintings ourselves.

2.1 Human Pose estimation

Human Pose Estimation (HPE) aims to detect human features from input data such as images and videos. It's an elementary part of computer vision with many applications among which are human action recognition (sign language), human tracking (surveillance), and human-computer interaction (video games). This is an extensively researched area with a diverse range of different techniques. This chapter will try to give an overview of all the many challenges and proposed solutions. The focus will be on deep learning models, which have surpassed classical solutions significantly. Specifically, around 2D monocular HPE eg., [12][3][13][14].

The human body has a high degree-of-freedom due to all the limbs, self-similar parts and body types, which may cause self-occlusion or rare/complex poses. The variations in configuration are made even larger due to clothing, lighting, foreground occlusion, as well as viewing angles and truncation, among others, as shown in fig. 2.1. This makes HPE one of the most difficult tasks in computer vision [15][2].

2.1.1 Representation

An important factor in HPE is how the pose will be represented. Depending on the needs of the problem you can have a skeleton-base, contour-base, or volume-base solution [2]2.2.

Skeleton-based model

The skeleton is build of a tree-structured set of keypoints that represent the joints of the human body. These can be explicitly described by their coordinates in 2D or 3D space [4]. More suitable for a Convolutional Neural Network (CNN) however is a heatmap which constructs a 2D Gaussian kernel around a keypoint [13][16]. They are easily implemented and became the dominant representation. While the skeleton-based model is a compact and flexible representation it suffers in this aspect by not being able to hold texture or shape information [3].

Contour representation

To capture the shape of the body parts, contour representation uses rectangles to estimate the body contours. These methods include cardboard models [17] and Active Shape Models (ASMs) [18] and were mainly in use in earlier HPE methods [2].

Volume representation

Volumetric geometric shapes can also be used as a method of representation. Earlier methods used simple shapes like cylinders, conics, and other shapes [19]. Volume representation is a 3D mesh that represents the human body. The most used model is Skinned Multi-Person Linear (SMPL), which includes natural pose-dependent deformations imitating soft-tissue dynamics [20].

For the purpose of our research, a simple model is the only thing we need. We only need to be aware of the most essential joints to label a pose. This makes the skeleton-based model the ideal representation to work with and will be the focus of further study.

2.1.2 Discriminative Methods and Generative Methods

Before deep learning became prominent in HPE there were already a number of different methods in use. Some of these methods are compatible with the deep learning methods and were thus adopted. An early distinction is between generative and discriminative methods.

Generative Model

A generative method will work with prior beliefs about the pose. More information about this can be found in the section about representation 2.1.1. It will project the pose on the image and verify it with the image data. If they don't comply, the pose is adjusted using the descent direction found by minimizing an error function [21].

Discriminative Model

Discriminative methods on the other hand, try to map the pose on the image data with learned models. There are several methods in this category, among which are the deep learning-based methods. The deep-learning methods are further categorized by the following sections.

2.1.3 Single-Person Methods

Single-person pose estimation will try to evaluate only one pose from an image. There are 2 major methods that are in use: regression methods and detection-based methods.

Regression-based Methods

The regression-based methods learn a network that maps all the body keypoints to the image-data directly as show in 2.3a.

The first successful deep learning model came from Toshev and Svegedy [4] and is considered the switch in paradigm from classic approaches to deep learning HPE. Toshev et al. uses a 7-layered model with 5 convolution layers and 2 fully-connected layers for the pose regressor, based on AlexNet for its simple but effective architecture [22]. They then cascade the resulting found keypoints of this model to itself where it refines it using the area around the keypoints. While the network is the same, the different stages will have different learned parameters. With every stage the found keypoints become more accurate.

Carreira et al. [23] introduce an Iterative Error Feedback which is a self-correcting model using top-down feedback. Using the image-data and a starting pose modeled as a heatmap, the model, based on GoogLeNet [24], will predict an error for each keypoint. The pose is then corrected based on the error and fed back into the model as a heatmap with the image. With each iteration it converges towards the solution instead of making the prediction in one go. Regression-based methods map the keypoints directly on the image, making it a non-linear problem. This will cause less robust generalization [13].

Heatmap/Detection-based Methods

The detection-based methods will first estimate the individual body parts using heatmaps, which leads to an easier optimization and a more robust generalization [14]. Most of the latest HPE methods use heatmaps because of this. After the joints are found they are then assembled to fit a human skeleton. This process is shown in 2.3b.

Tompson et al. [25] proposed a hybrid architecture where the detection of body parts is handled by a CNN and a Spatial-Model to bring those together. The first step produces many false-positives and these are removed in the second step by restricting joint inter-connectivity to enforce correct anatomy. They build on this in [26], where they used a cascade to refine predictions.

A fundamental work written by Wei et al. [6] combines convolution networks with Pose Machines [5]. Pose Machines is an iterative architecture which consists of 2 models: the first is used for stage 1 where it extracts potential heatmaps for the joints. The second model is used for subsequent stages where the result of the previous stage is fed in together with the results of its own convolution network on the input image. This gradually refines the predictions for the joints and their positioning. 2.5 shows this process.

Another influential work was being written at the same time by Newell et al. [7]. Similar to CPMs, this is also an iterative architecture. They suggest what they call a "stacked hourglass" network, where "hourglass" modules are repeated 2.6a. In an "hourglass" module, first, the features are downsampled and afterwards upsampled again 2.6b. This network captures different spatial relationships between joints at different resolutions. Several other works [27][**?**][28] have since improved on the network design.

Both these use intermediate supervision to tackle the problem of vanishing gradients. This still doesn't build a deep sub-network for feature extraction which limits the estimations. This has become less of a problem with the emergence of

Residual Network (ResNet)[29] which allows better back-propagation at deeper levels through shortcuts.

A more recent work by Sun et al. [8] maintains the high-resolution representations instead of working the high-resolution from the low-to-high sub-network. After a first high-resolution sub-network, it gradually adds high-to-low sub-networks in parallel to predict multi-resolution features. Before each branch, they apply multi-scale fusion, which joins the predicted features from each scale on each scale. Both are shown in 2.7. This network has proven very effective and inspired several variations [30][31][32].

With the emergence of neural networks also came Generative Adversarial Networks (GANs) [33], which proved useful for HPE. They are employed to improve constraints of joint inter-connectivity and infer occluded body parts.

Chen et al. [9] propose a structure-aware convolution network using a stacked hourglass as generator which generates heatmaps for each joint. They use 2 discriminators, one to discriminate between low- and high-confidence predictions, another for real and fake poses. The network is designed as a conditional Generative Adversarial Network (cGAN) [34], which allows it to generate pose heatmaps as well as occlusion heatmaps.

A more classic GAN is used by Chou et al. [35], where they use a stacked hourglass network for both the generator as the discriminator. The generator predicts the heatmaps for each joint and the discriminator distinguished between the real and fake ones.

2.1.4 Multi-Person Methods

With multi-person methods comes an extra layer of difficulty: they need to be able to detect each person separately. To solve this problem multi-person methods propose several solutions. The 2 most popular are top-down and bottom-up methods.

Top-Down Methods

This method will first try to detect all persons in the image with a human detector. Each person is cropped by the bounding box and a single-person estimator predicts a pose for each person.

Occlusion and truncation are a regular occurrence in multi-person scenes and inevitable problem. One of the early multi-person models, by Iqbal et al. [36], works towards creating a robust model against occlusion. It uses Faster RCCN [37] to detect the human boundaries. After which, it applies integer linear programming for each person's fully connected graph. This technique is similar to [38], but instead of working on all globally found joints it only considers local joints. It can also handle any kind of occlusion or truncation.

The use of a human detector comes with its own sort of problems. Fang et al. [39], with Regional Multi-person Pose Estimation (RPME), try to remedy these with 2 components: They try to tackle inaccurate bounding boxes with Symmetric Spatial Transformer Network, redundant detections with Parametric Pose Non-Maximum-Suppresion. They also propose a 3rd component, Pose-Guided Proposals Generator, which can augment training samples.

Papandreou et al. [40] use a 2 stage pipeline. In the first stage, they employ the Faster RCNN detector [37]. In the second stage, they estimate the pose in each found bounding box using their own network. It predicts heatmaps using a fully convolutional ResNet and use their own novel aggregation procedure. Afterwards, they do post-processing using keypoint-based Non-Maximum-Suppression (NMS) a method of their own making.

A continuous effort is taken by Chen et al. [9] to deal with occlusion and truncation. They suggest a 2 stage architecture, a Cascaded Pyramid Network (CPN) as seen in 2.8, where first the "simple" keypoints are captured with GlobalNet, a feature

pyramid network based on [41], and the "hard" keypoints are handled by their RefineNet, based on the upsampling and concatenating of HyperNet [42] and using an adapted stacked hourglass. They achieved great results and several others improved on their work [43][44].

In more recent research, a new method was become more powerful than CNNs. The Transformer [**?**], based on attention mechanisms which are used to optimize recurrent networks [45], eliminates the use of recurrent layers, keeping only the attention mechanisms. Yang et al. [46] use this architecture because allows for better understanding of the spatial dependencies and learns at a higher rate.

Bottom-Up Methods

A different approach is taken with bottom-up methods. They first locate all joints in the image and then assemble them in potential humans.

DeepCut by Pishchulin et al. [38], one of the first multi-person models using CNNs. Using Fast R-CNN [37], it detects the body parts and labels each. With the joints found, it then uses Integer Linear Programming (ILP) to assemble them. This method is very computationally expensive; NP-hard. Insafutdinov et al. [47] therefor introduce a stronger part detector and better optimization strategy with DeeperCut.

CPMs make a return with OpenPose by Cao et al. [48], they're used to predict the joints with heatmaps and Part Affinity Fields (PAFs). A part affinity field also encodes the position and orientation of the limb which makes the assembly of joints into different poses possible. They can achieve real-time results with this method, and several others have improved on their design [49][50][44]. The high performance is only applicable to high-resolution images. Low-resolution images or images with occlusions perform poorly.

Kreiss et al. [51] continue on the idea of fields and introduce the Part Intensity Fields (PIF) and Part Association Fields (PAF). First, they predict the location of the different joints with PIF. Afterwards, they use PAF to find the inter-joint relationships. They are able to outperform any previous OpenPose-based proposals on low-resolution and occlusions.

Newell et al. [52] introduce a new method called associative embedding for supervising CNNs both detection and grouping. This is a single-stage architecture as opposed to the two-staged architectures previously discussed. They make use of the stacked hourglass network from [7] with some small modifications.

Continuing on the idea of associative embedding, Cheng et al. [30] use HRNet [8] as backbone for their HigherHRNet. Their method focuses on the scale-variance problem; a problem which hasn't been studied much, so it can localize keypoints for small persons better. Lou et al. [53] introduce Scale-adaptive Heatmap Regression (SAHR) and Weight-adaptive Heatmap Regression (WAHR) to the scale-variance problem. SAHR adaptively adjusts the standard deviation of each heatmap corresponding with the scale of the person. WAHR rebalances the foreground and background samples, so SAHR can work to its fullest extent.

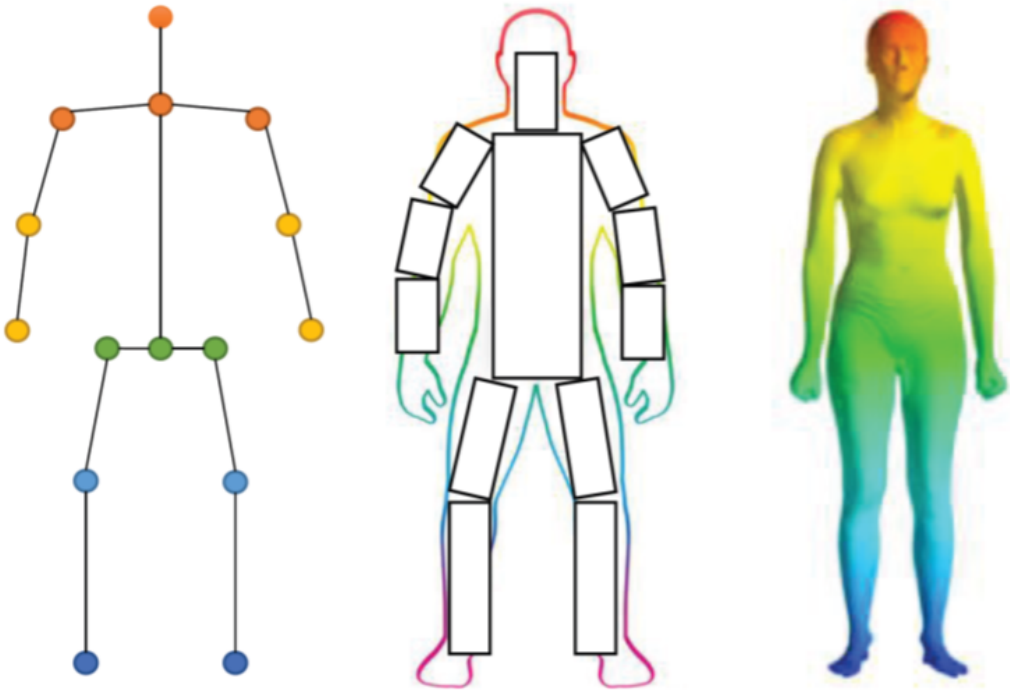
Summary

An important challenge for HPE is making predictions in scenes with hight occlusions. Top-down models achieve state-of-the art performance in almost all benchmark datasets [2]. Top-down models has difficulty with overlapping bodies and human detectors might fail finding humans there. To the same extent, bottom-up models will have greater inaccuracy with grouping in occluded scenes. Computationally, the top-down model's speed is limited by the number of people found. The

Figure 2.1: The various challenges HPE solutions face. Images from MPII dataset. [1][2]

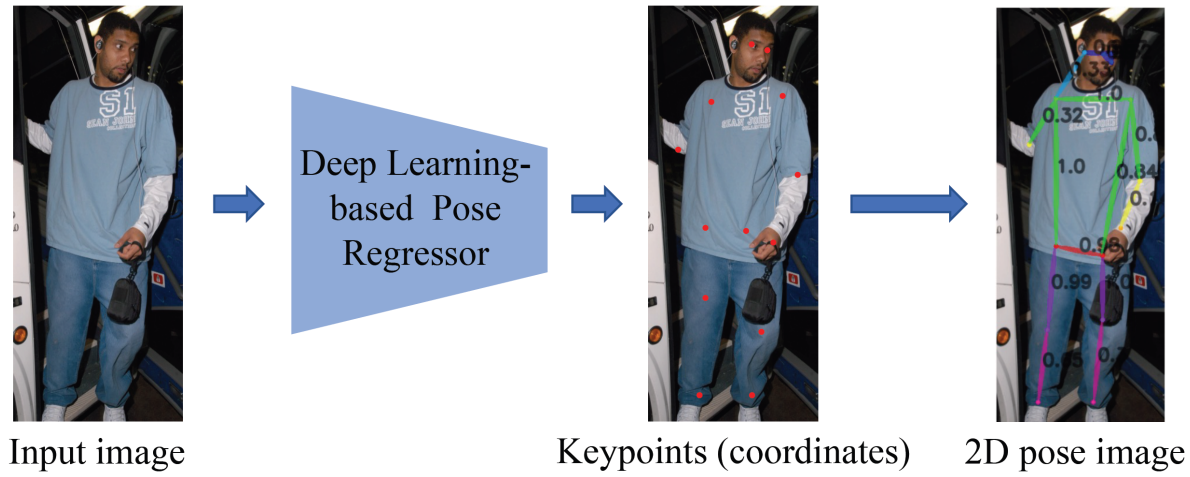
higher efficiency of bottom-up models, make them more suitable for real-time applications.

2.2 Image Style Transfer

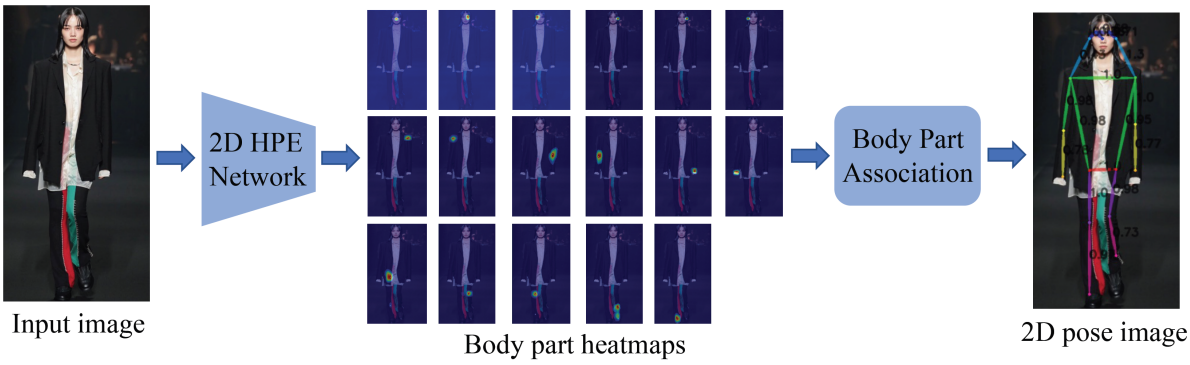


(a) Skeleton (b) Contour (c) Volume

Figure 2.2: Models for pose representation [3]

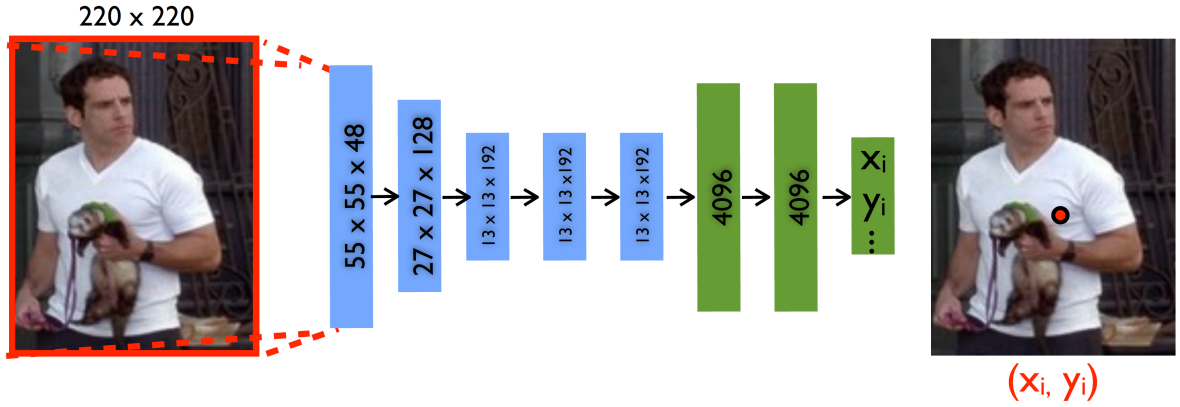


(a) Regression Methods

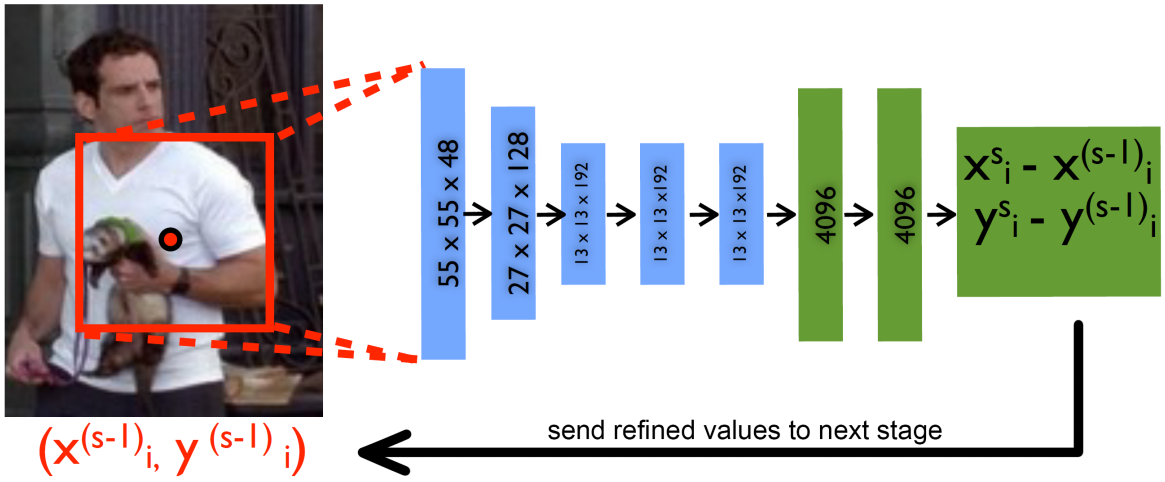


(b) Heatmap-based Methods

Figure 2.3: The different methods of single-person human pose estimation.[3]



(a) Initial stage



(b) Stage s

Figure 2.4: Convolution layers in blue and fully connected layers in green. The initial stage is applied to the whole images, while in stage s it will work on a sub-image based on the result of the previous stage.[4]

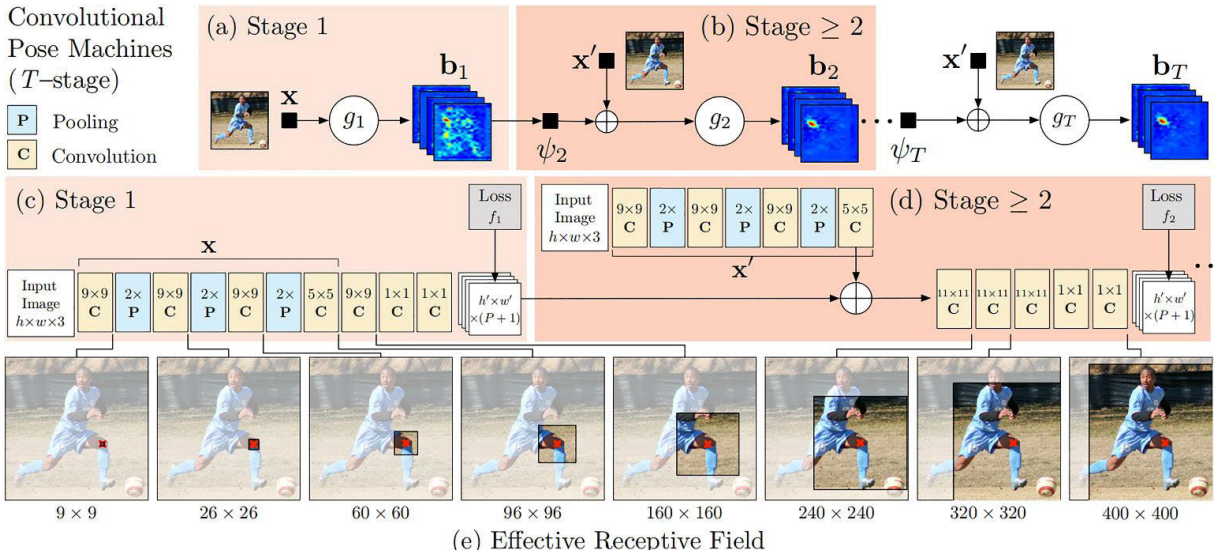
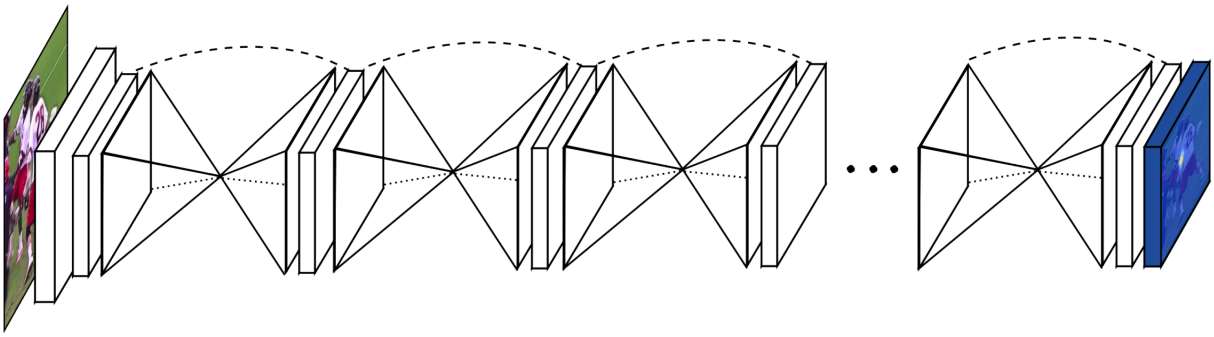
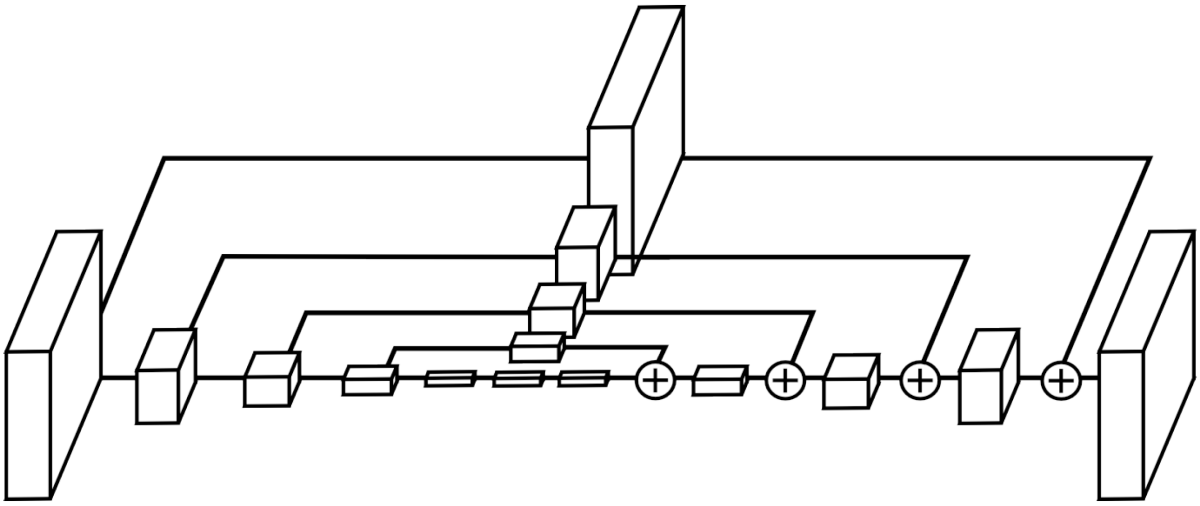


Figure 2.5: Architecture and receptive fields of CPMs. (a) and (b) represent the pose machine architecture.[5] (c) and (d) show the corresponding convolutional networks used by CPMs.[6]

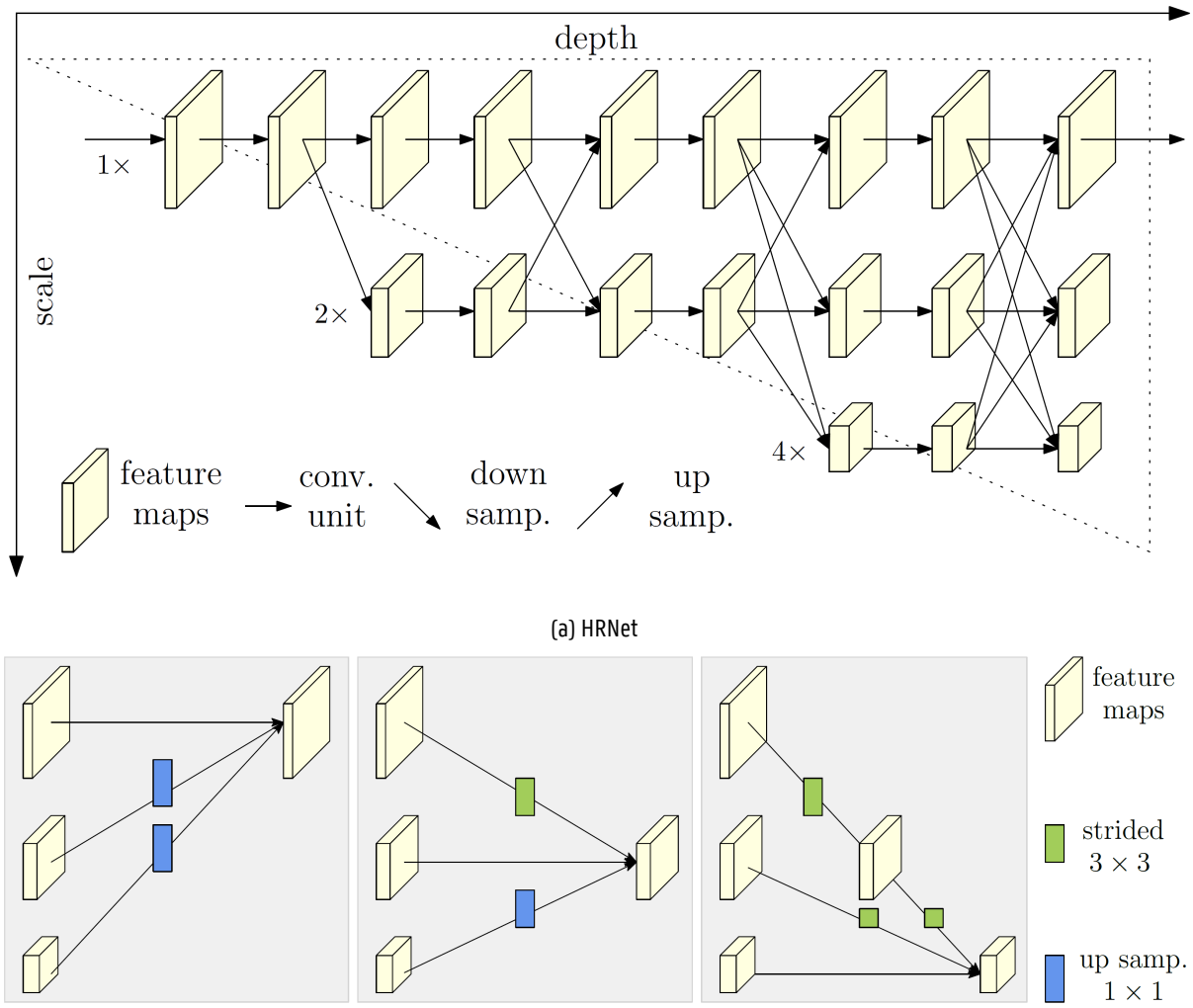


(a) Stacked Hourglass



(b) Hourglass Module

Figure 2.6: The structure of a "stacked hourglass" network and a single "hourglass" module.[7]



(a) HRNet

(b) Multi-Scale Fusion

Figure 2.7: The architecture of the High-Resolution network and how it applies multi-scale fusion.[8]

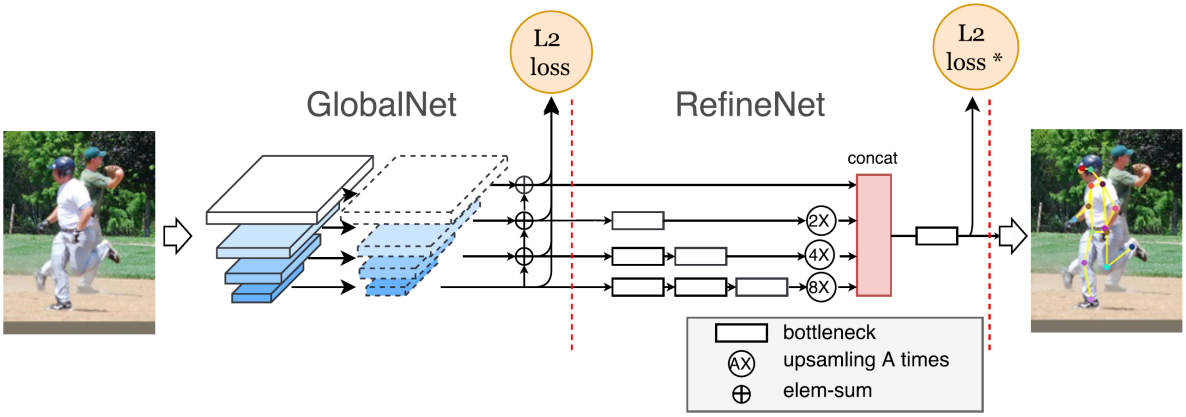


Figure 2.8: Cascaded Pyramid Network. "L2 loss*" means L2 loss with online hard keypoints mining.[9]

3

Titel tweede hoofdstuk

Vul aan.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

3.1 Sectie titel

Vul aan.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent

vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo. Voorbeeld figuur.

4

Titel derde hoofdstuk

Vul aan.

Nulla ac nisl. Nullam urna nulla, ullamcorper in, interdum sit amet, gravida ut, risus. Aenean ac enim. In luctus. Phasellus eu quam vitae turpis viverra pellentesque. Duis feugiat felis ut enim. Phasellus pharetra, sem id porttitor sodales, magna nunc aliquet nibh, nec blandit nisl mauris at pede. Suspendisse risus risus, lobortis eget, semper at, imperdiet sit amet, quam. Quisque scelerisque dapibus nibh. Nam enim. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Nunc ut metus. Ut metus justo, auctor at, ultrices eu, sagittis ut, purus. Aliquam aliquam.

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetuer quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Etiam suscipit aliquam arcu. Aliquam sit amet est ac purus bibendum congue. Sed in eros. Morbi non orci. Pellentesque mattis lacinia elit. Fusce molestie velit in ligula. Nullam et orci vitae nibh vulputate auctor. Aliquam eget purus. Nulla auctor wisi sed ipsum. Morbi porttitor tellus ac enim. Fusce ornare. Proin ipsum enim, tincidunt in, ornare venenatis, molestie a, augue. Donec vel pede in lacus sagittis porta. Sed hendrerit ipsum quis nisl. Suspendisse quis massa ac nibh pretium cursus. Sed sodales. Nam eu neque quis pede dignissim ornare. Maecenas eu purus ac urna tincidunt congue.

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Maecenas non massa. Vestibulum pharetra nulla at lorem. Duis quis quam id lacus dapibus interdum. Nulla lorem. Donec ut ante quis dolor bibendum condimentum. Etiam egestas tortor vitae lacus. Praesent cursus. Mauris bibendum pede at elit. Morbi et felis a lectus interdum facilisis. Sed suscipit gravida turpis. Nulla at lectus. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Praesent nonummy luctus nibh. Proin turpis nunc, congue eu, egestas ut, fringilla at, tellus. In hac habitasse platea dictumst.

4.1 Sectie titel

Vul aan.

Phasellus id magna. Duis malesuada interdum arcu. Integer metus. Morbi pulvinar pellentesque mi. Suspendisse sed est eu magna molestie egestas. Quisque mi lorem, pulvinar eget, egestas quis, luctus at, ante. Proin auctor vehicula purus. Fusce ac nisl aliquam ante hendrerit pellentesque. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Morbi wisi. Etiam arcu mauris, facilisis sed, eleifend non, nonummy ut, pede. Cras ut lacus tempor metus mollis placerat. Vivamus eu tortor vel metus interdum malesuada.

Sed eleifend, eros sit amet faucibus elementum, urna sapien consectetuer mauris, quis egestas leo justo non risus. Morbi non felis ac libero vulputate fringilla. Mauris libero eros, lacinia non, sodales quis, dapibus porttitor, pede. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Morbi dapibus mauris condimentum nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Etiam sit amet erat. Nulla varius. Etiam tincidunt dui vitae turpis. Donec leo. Morbi vulputate convallis est. Integer aliquet. Pellentesque aliquet sodales urna.

Nullam eleifend justo in nisl. In hac habitasse platea dictumst. Morbi nonummy. Aliquam ut felis. In velit leo, dictum vitae, posuere id, vulputate nec, ante. Maecenas vitae pede nec dui dignissim suscipit. Morbi magna. Vestibulum id purus eget velit laoreet laoreet. Praesent sed leo vel nibh convallis blandit. Ut rutrum. Donec nibh. Donec interdum. Fusce sed pede sit amet elit rhoncus ultrices. Nullam at enim vitae pede vehicula iaculis.

4.2 Sectie titel2

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Voorbeeld figuur.

Subtitel

Vul aan

4.2.1 Functie is_isbn

Voorbeeld listing.

```
def is_isbn(isbn: str) -> bool:
    if not type(isbn) is str:
        return False
    if len(isbn) != 10:
        return False
    if not isbn[:9].isdigit():
        return False
    if not isbn[9] in "0123456789X":
        return False
    som = 0
    for i in range(9):
        cijfer = int(isbn[i])
        som += (i + 1) * cijfer
    if isbn[9] == "X":
        laatste_cijfer = 10
    else:
        laatste_cijfer = int(isbn[9])
    return laatste_cijfer == som % 11
```

Code Fragment 4.1: Functie is_isbn

Conclusie

Vul aan.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Ethische en maatschappelijke reflectie

Vul aan.

Meer informatie kan je opzoeken op https://www.sdgs.be/nl/sdgs

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Referenties

[1] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2d human pose estimation: New benchmark and state of the art analysis," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3686–3693.

[2] Y. Chen, Y. Tian, and M. He, "Monocular human pose estimation: A survey of deep learning-based methods," *CoRR*, vol. abs/2006.01423, 2020. [Online]. Available: https://arxiv.org/abs/2006.01423

[3] C. Zheng, W. Wu, T. Yang, S. Zhu, C. Chen, R. Liu, J. Shen, N. Kehtarnavaz, and M. Shah, "Deep learning-based human pose estimation: A survey," *CoRR*, vol. abs/2012.13392, 2020. [Online]. Available: https://arxiv.org/abs/2012.13392

[4] A. Toshev and C. Szegedy, "DeepPose: Human pose estimation via deep neural networks," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, jun 2014. [Online]. Available: https://doi.org/10.1109%2Fcvpr.2014.214

[5] V. Ramakrishna, D. Munoz, M. Hebert, J. Andrew Bagnell, and Y. Sheikh, "Pose machines: Articulated pose estimation via inference machines," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 33–47.

[6] S. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," *CoRR*, vol. abs/1602.00134, 2016. [Online]. Available: http://arxiv.org/abs/1602.00134

[7] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," *CoRR*, vol. abs/1603.06937, 2016. [Online]. Available: http://arxiv.org/abs/1603.06937

[8] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," *CoRR*, vol. abs/1902.09212, 2019. [Online]. Available: http://arxiv.org/abs/1902.09212

[9] Y. Chen, C. Shen, X. Wei, L. Liu, and J. Yang, "Adversarial posenet: A structure-aware convolutional network for human pose estimation," *CoRR*, vol. abs/1705.00389, 2017. [Online]. Available: http://arxiv.org/abs/1705.00389

[10] M.-C. Marinescu, A. Reshetnikov, and J. M. López, "Improving object detection in paintings based on time contexts," in *2020 International Conference on Data Mining Workshops (ICDMW)*, 2020, pp. 926–932.

[11] M. Sabatelli, N. Banar, M. Cocriamont, E. Coudyzer, K. Lasaracina, W. Daelemans, P. Geurts, and M. Kestemont, "Advances in digital music iconography: Benchmarking the detection of musical instruments in unrestricted, non-photorealistic images from the artistic domain," *Digital Humanities Quarterly*, vol. 15, no. 1, February 2021.

[12] T. L. Munea, Y. Z. Jembre, H. T. Weldegebriel, L. Chen, C. Huang, and C. Yang, "The progress of human pose estimation: A survey and taxonomy of models applied in 2d human pose estimation," *IEEE Access*, vol. 8, pp. 133 330–133 348, 2020.

[13] W. Liu, Q. Bao, Y. Sun, and T. Mei, "Recent advances in monocular 2d and 3d human pose estimation: A deep learning perspective," *CoRR*, vol. abs/2104.11536, 2021. [Online]. Available: https://arxiv.org/abs/2104.11536

[14] H. Chen, R. Feng, S. Wu, H. Xu, F. Zhou, and Z. Liu, "2d human pose estimation: a survey," *Multimedia Systems*, pp. 1–24, 2022.

[15] A. Jain, J. Tompson, M. Andriluka, G. W. Taylor, and C. Bregler, "Learning human pose estimation features with convolutional networks," 2014.

[16] Z. Luo, Z. Wang, Y. Huang, T. Tan, and E. Zhou, "Rethinking the heatmap regression for bottom-up human pose estimation," *CoRR*, vol. abs/2012.15175, 2020. [Online]. Available: https://arxiv.org/abs/2012.15175

[17] S. Ju, M. Black, and Y. Yacoob, "Cardboard people: a parameterized model of articulated image motion," in *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, 1996, pp. 38–44.

[18] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1077314285710041

[19] H. Sidenbladh, F. De la Torre, and M. Black, "A framework for modeling the appearance of 3d articulated figures," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, 2000, pp. 368–375.

[20] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. Black, "Smpl: a skinned multi-person linear model," vol. 34, 11 2015.

[21] G. Pons-Moll and B. Rosenhahn, *Model-Based Pose Estimation*. London: Springer London, 2011, pp. 139–170. [Online]. Available: https://doi.org/10.1007/978-0-85729-997-0_9

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[23] J. Carreira, P. Agrawal, K. Fragkiadaki, and J. Malik, "Human pose estimation with iterative error feedback," *CoRR*, vol. abs/1507.06550, 2015. [Online]. Available: http://arxiv.org/abs/1507.06550

[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *CoRR*, vol. abs/1409.4842, 2014. [Online]. Available: http://arxiv.org/abs/1409.4842

[25] J. Tompson, A. Jain, Y. LeCun, and C. Bregler, "Joint training of a convolutional network and a graphical model for human pose estimation," *CoRR*, vol. abs/1406.2984, 2014. [Online]. Available: http://arxiv.org/abs/1406.2984

[26] J. Tompson, R. Goroshin, A. Jain, Y. LeCun, and C. Bregler, "Efficient object localization using convolutional networks," *CoRR*, vol. abs/1411.4280, 2014. [Online]. Available: http://arxiv.org/abs/1411.4280

[27] W. Yang, S. Li, W. Ouyang, H. Li, and X. Wang, "Learning feature pyramids for human pose estimation," *CoRR*, vol. abs/1708.01101, 2017. [Online]. Available: http://arxiv.org/abs/1708.01101

[28] C. Chou, J. Chien, and H. Chen, "Self adversarial training for human pose estimation," *CoRR*, vol. abs/1707.02439, 2017. [Online]. Available: http://arxiv.org/abs/1707.02439

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[30] B. Cheng, B. Xiao, J. Wang, H. Shi, T. S. Huang, and L. Zhang, "Bottom-up higher-resolution networks for multi-person pose estimation," *CoRR*, vol. abs/1908.10357, 2019. [Online]. Available: http://arxiv.org/abs/1908.10357

[31] C. Yu, B. Xiao, C. Gao, L. Yuan, L. Zhang, N. Sang, and J. Wang, "Lite-hrnet: A lightweight high-resolution network," *CoRR*, vol. abs/2104.06403, 2021. [Online]. Available: https://arxiv.org/abs/2104.06403

[32] Y. Yuan, R. Fu, L. Huang, W. Lin, C. Zhang, X. Chen, and J. Wang, "Hrformer: High-resolution transformer for dense prediction," *CoRR*, vol. abs/2110.09408, 2021. [Online]. Available: https://arxiv.org/abs/2110.09408

[33] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[34] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *CoRR*, vol. abs/1411.1784, 2014. [Online]. Available: http://arxiv.org/abs/1411.1784

[35] C. Chou, J. Chien, and H. Chen, "Self adversarial training for human pose estimation," *CoRR*, vol. abs/1707.02439, 2017. [Online]. Available: http://arxiv.org/abs/1707.02439

[36] U. Iqbal and J. Gall, "Multi-person pose estimation with local joint-to-person associations," *CoRR*, vol. abs/1608.08526, 2016. [Online]. Available: http://arxiv.org/abs/1608.08526

[37] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *CoRR*, vol. abs/1506.01497, 2015. [Online]. Available: http://arxiv.org/abs/1506.01497

[38] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. V. Gehler, and B. Schiele, "Deepcut: Joint subset partition and labeling for multi person pose estimation," *CoRR*, vol. abs/1511.06645, 2015. [Online]. Available: http://arxiv.org/abs/1511.06645

[39] H. Fang, S. Xie, and C. Lu, "RMPE: regional multi-person pose estimation," *CoRR*, vol. abs/1612.00137, 2016. [Online]. Available: http://arxiv.org/abs/1612.00137

[40] G. Papandreou, T. Zhu, N. Kanazawa, A. Toshev, J. Tompson, C. Bregler, and K. P. Murphy, "Towards accurate multi-person pose estimation in the wild," *CoRR*, vol. abs/1701.01779, 2017. [Online]. Available: http://arxiv.org/abs/1701.01779

[41] T. Lin, P. Dollár, R. B. Girshick, K. He, B. Hariharan, and S. J. Belongie, "Feature pyramid networks for object detection," *CoRR*, vol. abs/1612.03144, 2016. [Online]. Available: http://arxiv.org/abs/1612.03144

[42] T. Kong, A. Yao, Y. Chen, and F. Sun, "Hypernet: Towards accurate region proposal generation and joint object detection," *CoRR*, vol. abs/1604.00600, 2016. [Online]. Available: http://arxiv.org/abs/1604.00600

[43] K. Su, D. Yu, Z. Xu, X. Geng, and C. Wang, "Multi-person pose estimation with enhanced channel-wise and spatial information," *CoRR*, vol. abs/1905.03466, 2019. [Online]. Available: http://arxiv.org/abs/1905.03466

[44] W. Li, Z. Wang, B. Yin, Q. Peng, Y. Du, T. Xiao, G. Yu, H. Lu, Y. Wei, and J. Sun, "Rethinking on multi-stage networks for human pose estimation," *CoRR*, vol. abs/1901.00148, 2019. [Online]. Available: http://arxiv.org/abs/1901.00148

[45] Y. Kim, C. Denton, L. Hoang, and A. M. Rush, "Structured attention networks," *CoRR*, vol. abs/1702.00887, 2017. [Online]. Available: http://arxiv.org/abs/1702.00887

[46] S. Yang, Z. Quan, M. Nie, and W. Yang, "Transpose: Towards explainable human pose estimation by transformer," *CoRR*, vol. abs/2012.14214, 2020. [Online]. Available: https://arxiv.org/abs/2012.14214

[47] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele, "Deepercut: A deeper, stronger, and faster multi-person pose estimation model," *CoRR*, vol. abs/1605.03170, 2016. [Online]. Available: http://arxiv.org/abs/1605.03170

[48] Z. Cao, T. Simon, S. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," *CoRR*, vol. abs/1611.08050, 2016. [Online]. Available: http://arxiv.org/abs/1611.08050

[49] X. Zhu and Y. Jiang, "Multi-person pose estimation for posetrack with enhanced part affinity fields," 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:52563463

[50] G. Hidalgo, Y. Raaj, H. Idrees, D. Xiang, H. Joo, T. Simon, and Y. Sheikh, "Single-network whole-body pose estimation," *CoRR*, vol. abs/1909.13423, 2019. [Online]. Available: http://arxiv.org/abs/1909.13423

[51] S. Kreiss, L. Bertoni, and A. Alahi, "Pifpaf: Composite fields for human pose estimation," *CoRR*, vol. abs/1903.06593, 2019. [Online]. Available: http://arxiv.org/abs/1903.06593

[52] A. Newell and J. Deng, "Associative embedding: End-to-end learning for joint detection and grouping," *CoRR*, vol. abs/1611.05424, 2016. [Online]. Available: http://arxiv.org/abs/1611.05424

[53] Z. Luo, Z. Wang, Y. Huang, T. Tan, and E. Zhou, "Rethinking the heatmap regression for bottom-up human pose estimation," *CoRR*, vol. abs/2012.15175, 2020. [Online]. Available: https://arxiv.org/abs/2012.15175

Bijlagen

Bijlage 1

Toelichting bijlage.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Bijlage 2

Toelichting bijlage.

Nulla mattis luctus nulla. Duis commodo velit at leo. Aliquam vulputate magna et leo. Nam vestibulum ullamcorper leo. Vestibulum condimentum rutrum mauris. Donec id mauris. Morbi molestie justo et pede. Vivamus eget turpis sed nisl cursus tempor. Curabitur mollis sapien condimentum nunc. In wisi nisl, malesuada at, dignissim sit amet, lobortis in, odio. Aenean consequat arcu a ante. Pellentesque porta elit sit amet orci. Etiam at turpis nec elit ultricies imperdiet. Nulla facilisi. In hac habitasse platea dictumst. Suspendisse viverra aliquam risus. Nullam pede justo, molestie nonummy, scelerisque eu, facilisis vel, arcu.

Curabitur tellus magna, porttitor a, commodo a, commodo in, tortor. Donec interdum. Praesent scelerisque. Maecenas posuere sodales odio. Vivamus metus lacus, varius quis, imperdiet quis, rhoncus a, turpis. Etiam ligula arcu, elementum a, venenatis quis, sollicitudin sed, metus. Donec nunc pede, tincidunt in, venenatis vitae, faucibus vel, nibh. Pellentesque wisi. Nullam malesuada. Morbi ut tellus ut pede tincidunt porta. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam congue neque id dolor.

Donec et nisl at wisi luctus bibendum. Nam interdum tellus ac libero. Sed sem justo, laoreet vitae, fringilla at, adipiscing ut, nibh. Maecenas non sem quis tortor eleifend fermentum. Etiam id tortor ac mauris porta vulputate. Integer porta neque vitae massa. Maecenas tempus libero a libero posuere dictum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aenean quis mauris sed elit commodo placerat. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Vivamus rhoncus tincidunt libero. Etiam elementum pretium justo. Vivamus est. Morbi a tellus eget pede tristique commodo. Nulla nisl. Vestibulum sed nisl eu sapien cursus rutrum.

Nulla non mauris vitae wisi posuere convallis. Sed eu nulla nec eros scelerisque pharetra. Nullam varius. Etiam dignissim elementum metus. Vestibulum faucibus, metus sit amet mattis rhoncus, sapien dui laoreet odio, nec ultricies nibh augue a enim. Fusce in ligula. Quisque at magna et nulla commodo consequat. Proin accumsan imperdiet sem. Nunc porta. Donec feugiat mi at justo. Phasellus facilisis ipsum quis ante. In ac elit eget ipsum pharetra faucibus. Maecenas viverra nulla in massa.

Nulla ac nisl. Nullam urna nulla, ullamcorper in, interdum sit amet, gravida ut, risus. Aenean ac enim. In luctus. Phasellus eu quam vitae turpis viverra pellentesque. Duis feugiat felis ut enim. Phasellus pharetra, sem id porttitor sodales, magna nunc aliquet nibh, nec blandit nisl mauris at pede. Suspendisse risus risus, lobortis eget, semper at, imperdiet sit amet, quam. Quisque scelerisque dapibus nibh. Nam enim. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Nunc ut metus. Ut metus justo, auctor at, ultrices eu, sagittis ut, purus. Aliquam aliquam.

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetuer quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Etiam suscipit aliquam arcu. Aliquam sit amet est ac purus bibendum congue. Sed in eros. Morbi non orci. Pellentesque mattis lacinia elit. Fusce molestie velit in ligula. Nullam et orci vitae nibh vulputate auctor. Aliquam eget purus. Nulla auctor wisi sed ipsum. Morbi porttitor tellus ac enim. Fusce ornare. Proin ipsum enim, tincidunt in, ornare venenatis, molestie a,

augue. Donec vel pede in lacus sagittis porta. Sed hendrerit ipsum quis nisl. Suspendisse quis massa ac nibh pretium cursus. Sed sodales. Nam eu neque quis pede dignissim ornare. Maecenas eu purus ac urna tincidunt congue.

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Maecenas non massa. Vestibulum pharetra nulla at lorem. Duis quis quam id lacus dapibus interdum. Nulla lorem. Donec ut ante quis dolor bibendum condimentum. Etiam egestas tortor vitae lacus. Praesent cursus. Mauris bibendum pede at elit. Morbi et felis a lectus interdum facilisis. Sed suscipit gravida turpis. Nulla at lectus. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Praesent nonummy luctus nibh. Proin turpis nunc, congue eu, egestas ut, fringilla at, tellus. In hac habitasse platea dictumst.

Vivamus eu tellus sed tellus consequat suscipit. Nam orci orci, malesuada id, gravida nec, ultricies vitae, erat. Donec risus turpis, luctus sit amet, interdum quis, porta sed, ipsum. Suspendisse condimentum, tortor at egestas posuere, neque metus tempor orci, et tincidunt urna nunc a purus. Sed facilisis blandit tellus. Nunc risus sem, suscipit nec, eleifend quis, cursus quis, libero. Curabitur et dolor. Sed vitae sem. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Maecenas ante. Duis ullamcorper enim. Donec tristique enim eu leo. Nullam molestie elit eu dolor. Nullam bibendum, turpis vitae tristique gravida, quam sapien tempor lectus, quis pretium tellus purus ac quam. Nulla facilisi.

Duis aliquet dui in est. Donec eget est. Nunc lectus odio, varius at, fermentum in, accumsan non, enim. Aliquam erat volutpat. Proin sit amet nulla ut eros consectetuer cursus. Phasellus dapibus aliquam justo. Nunc laoreet. Donec consequat placerat magna. Duis pretium tincidunt justo. Sed sollicitudin vestibulum quam. Nam quis ligula. Vivamus at metus. Etiam imperdiet imperdiet pede. Aenean turpis. Fusce augue velit, scelerisque sollicitudin, dictum vitae, tempor et, pede. Donec wisi sapien, feugiat in, fermentum ut, sollicitudin adipiscing, metus.

Donec vel nibh ut felis consectetuer laoreet. Donec pede. Sed id quam id wisi laoreet suscipit. Nulla lectus dolor, aliquam ac, fringilla eget, mollis ut, orci. In pellentesque justo in ligula. Maecenas turpis. Donec eleifend leo at felis tincidunt consequat. Aenean turpis metus, malesuada sed, condimentum sit amet, auctor a, wisi. Pellentesque sapien elit, bibendum ac, posuere et, congue eu, felis. Vestibulum mattis libero quis metus scelerisque ultrices. Sed purus.

Donec molestie, magna ut luctus ultrices, tellus arcu nonummy velit, sit amet pulvinar elit justo et mauris. In pede. Maecenas euismod elit eu erat. Aliquam augue wisi, facilisis congue, suscipit in, adipiscing et, ante. In justo. Cras lobortis neque ac ipsum. Nunc fermentum massa at ante. Donec orci tortor, egestas sit amet, ultrices eget, venenatis eget, mi. Maecenas vehicula leo semper est. Mauris vel metus. Aliquam erat volutpat. In rhoncus sapien ac tellus. Pellentesque ligula.