

Technology for Translating, Visualizing, and Generating Recipes

Multimodal Machine Translation of
Chinese into English

Chan Sin-wai

First published 2026

ISBN: 9781032308425 (hbk)

ISBN: 9781032308432 (pbk)

ISBN: 9781003306948 (ebk)

2 *TransRecipe* for Translating Recipes through Machine Processing

CC BY-NC-ND

DOI: 10.4324/9781003306948-3



Routledge
Taylor & Francis Group
LONDON AND NEW YORK

2 *TransRecipe* for Translating Recipes through Machine Processing

Introduction

Machine processing of recipes started many years ago. *TransRecipe* was one such effort made a quarter of a century ago by the author for a cookbook translation. This pioneering attempt to explore the suitability of a machine translation system to produce quality translations for specific domains must be seen against the broader context of the development of computer science and the growth of translation studies. Since the 1970s and 1980s, when translation studies came to maturity as an independent academic discipline, it has firmly positioned itself in the midst of the technological revolution. With communication becoming faster and more efficient, translation also has to be created and delivered rapidly, accurately, and in formats that fit the various modalities now available. These modalities pose new challenges for translation, and exert an enormous impact on the discipline development on practical, theoretical, and educational levels.

In fact, the importance of machine translation and machine-aided translation in this digital age cannot be over-emphasized. One of the primary functions of machine translation, as we know, is to serve as a translation aid and remove linguistic barriers in order to cope with an ever-increasing volume of materials that appear on the Internet in other languages that we do not know. While the Internet has revolutionized our way of life, the world has become a global village where a multitude of languages are used. Machine translation has made global communication possible. This is aptly pointed out by Colin Haynes (1998: 3): “For the first time in history, we now have the capability to communicate with each other across the boundaries of time, space and language.”

On a more microscopic level, machine translation is also important in regional development. With China’s entry into the World Trade Organisation and to preserve Hong Kong’s role as the world’s leading financial centre, the demand for competent translators in this part of the world has been enormous. The shortage of human translators will be minimized by the supply of reliable domain-specific machine translation systems. This can be illustrated by a comparison between the amount of translation that can be produced by

a human translator with that of a machine translation system. It is estimated that a human translator can, on average, translate 2,000 words per day on an eight-hour working basis, but the speed of an average translation system is around 3,000 words per second. In eight hours, a system can translate as many as 86,400,000 words, which is 43,200 times faster than a human translator. The current systems on the market are much faster than before.

On the other hand, the quality and marketability of machine translation systems have been improved considerably. Machine translation software is becoming more and more viable, affordable, and user-friendly. The production of low-cost and easy-to-use software, coupled with the growing need to translate online information, means that machine translation systems are now widely used and web translation is a leading trend. In fact, machine translation software is more than an academic pursuit, it is also a profitable business. According to Muriel Vasconcellos (1996: 1), “today, there are more than 500 vendors of machine translation software for the personal computer around the world, and among them they put out well over 1,000 products. One of the vendors, Globalink, sells its extensive line of software in at least 6,000 stores in North America alone, and at present Europe is its fastest-growing market.” Obviously, machine translation is no longer a matter that is only of interest to academics, such as university linguists, computer engineers, and translators, but is also an area from which commercial interests can be generated.

Despite this, machine translation systems in the market before the introduction of neural machine translation are far from satisfactory. One of the major reasons that accounts for the failure of machine translation is the over-dependence on the application of linguistic concepts to the construction of a machine translation system. The linguistic theories that have been used so far have not produced very effective and reliable systems. We certainly need to explore new areas and new methodologies to design new systems that work, and basically, this has a lot to do with the distinction between what is known as a “language system” and “language in use”. What the theorists have focused on are formalistic analyses of the grammatical and linguistic features of the languages concerned, to the neglect of empirical variations in actual usage within specific contexts. This accounts for the success of “practical systems” based on domain-specific grammars, and why a linguistics-oriented system with a complicated flowchart filled with technical jargon has not yielded results which are translationally acceptable.

A fully automatic high-quality machine translation system for general applications for any specific language pair still eludes us. For professional translators, machine-aided translation in the form of an interactive translation workstation will be extremely useful and is expected to become very popular in the near future. There is in fact no rivalry between machine translation and human translation because they complement, rather than compete with each other. The human translator can help to improve the quality of machine output, and machine translation can speed up the process of human translation. Certain routine documents can be handled by the machine, while the

human translator can spend more time on the production of quality translation of literary, legal, and diplomatic documents.

Another point worth noting is that the design of a machine translation system has so far been in the hands of computer scientists and linguists. The views of translators, who are the professionals as well as the users, have not been considered in the construction of machine translation software. It is about time that human translators play a more active role in machine translation. And it can be expected that the role of translators in the design of a machine translation system will become increasingly important. A translational approach to machine translation will be a new development in the field. It can also be expected that a “translation-oriented” system with substantial input from practising translators in the areas of translation process, translation methodology and translation generation, will soon be in place.

TransRecipe is a fully automatic translation system for translating Chinese cookbooks into English based on the transfer model which combines the corpus-based, example-based, pattern-based, and rule-based approaches. With a lexical database of around 20,000 entries relating to Chinese food and drink, a general dictionary of around 2,000 frequently-used expressions, a database of 200 examples, and a total of 700 global parsing rules, the system can automatically translate a recipe from Chinese into English within seconds without any pre-processing or post-editing.

This system sets out to achieve several goals in the field of machine translation which have not yet been realized by other Chinese-English translation systems on the market. First, it is the first fully-automatic translation system specially designed for the rendition of Chinese recipes into English. Second, it applies a large number of translation methods in the parsing stage in order to produce a high-quality output that is translationally acceptable. Third, it serves as a pilot system to illustrate the mechanical translation of procedural texts, which can have wider applications to other types of texts, such as instruction manuals, financial documents, and scientific writings.

The making of *TransRecipe* also serves to highlight the importance of using translation methods in the construction of a machine translation system to produce a good translation. This particular orientation is labelled as a “translational approach” to distinguish it from approaches based largely on linguistic and computational concepts. It is undeniable that the various machine systems for translating between Chinese and English have been designed mainly by scholars and specialists in the fields of computer science and linguistics, and rarely by translation experts. As a result, most of the outputs are at best linguistically equivalent texts hardly intelligible to the reader. Also, as we all know, machine translation works at the syntactical level, and this means that methods frequently used in translating sentences to produce effective translations have not been put to good use in machine translation. Methods of breaking down long sentences into shorter ones, for example, would certainly make it much easier for a machine translation system to produce a more accurate target text.

8 *Technology for Translating, Visualizing, and Generating Recipes*

It is hoped that *TransRecipe* will help the English-reading public in different parts of the world to practise the Chinese culinary art, which is a very important part of Chinese culture, and it will also show that the experience of a human translator can make a great deal of difference to the quality of machine translation.

Recipes: Grammatical Analyses

One of the most important steps in setting up a translation system is to analyse the source text linguistically, or more specifically, grammatically, since the sentence is the linguistic unit for a machine system. What is grammar? According to *Webster's Third New International Dictionary of the English Language* (Gove, 1986: 986), grammar is defined as:

A branch of linguistic study that deals with the classes of words, their inflections or other means of indicating relation to each other, and their functions and relations in the sentence as employed according to established usage and that is sometimes extended to include related matter such as phonology, prosody, language history, orthography, orthoepy, etymology, or semantics.

According to one of the most influential works on English grammar, grammar is to “include both syntax and that aspect of morphology (the internal structure of words) that deals with inflections (or accidence). (Quirk et al., 1985: 12). A lively description of grammar is given by Peter Newmark (1988: 125) when he relates grammar to translation:

Grammar is the skeleton of a text... Grammar gives you the general and main facts about a text: statements, questions, requests, purpose, reason, condition, time, place, doubt, feeling, certainty. Grammar indicates who does what to whom, why, where, when, how.

Translating, we must admit, is more than grammatical transfer. But it should also be emphasised that grammar can be used to break down the source language sentence for meaning acquisition, and to verify the semantic correctness and syntactic structure of the target text. This explains why traditional grammar is used in this machine translation system. Let us discuss traditional grammar using different parts of speech (except for interjection which is never used in recipes).

(a) Adjectives

An adjective is a word that is used to modify a noun, expressing a characteristic quality or attribute. English adjectives can be both attributive and predicative,

premodified by intensifiers and can take comparatives and superlatives. In recipes, one of the most commonly used adjectives is *shaoxu* 少許, the translation of which depends on its collocates. The following equivalents have been used when translating *shaoxu*:

麻油少許	<i>a dash of</i> sesame oil
荔枝醋少許	<i>a little</i> lychee vinegar
芫荽少許	<i>a pick of</i> parsley
胡椒粉少許	<i>a pinch of</i> pepper
鷹粟粉少許	<i>some</i> cornflour

Another point to note is that very often, adjectives in recipes are not translated according to the dictionary meanings. *Liang* 靚, for example, is usually translated as “beautiful”. In Chinese recipes, it is translated as “fresh”, such as translating *liang shore* 靚豬肉 into “fresh pork”.

(b) Adverbs

An adverb is a word or phrase whose main function is to modify a verb, an adjective, a sentence, or another adverb. There are several types of adverbs: adverb of degree, of direction, of frequency, of manner, of place, of recurrence, of sequence, and of time. Among them, the following two are relevant to recipes.

Adverb of Manner

This type of adverb, such as carefully, easily, quickly, and slowly, refers to the manner something is or has been done. In the *TransRecipe* system, some manner adverbs in recipes are usually not translated, such as the word *tong* 同, meaning “together”.

鹹豬肉、豬腿肉同飛水。

Scald salted pork and pork from hind leg in boiling water.

Adverb of Time

An adverb of time, also known as a temporal adverb, refers to the time something was or will be done. Adverbs such as now, hourly, yesterday are adverbs of time. Recipes inevitably have such kinds of adverbs as an indication of duration has to be included in cookbooks.

加入調味料蚊十分鐘，加入冬菰再炊五分鐘。

Put in seasoning and braise for 10 minutes. Add black mushrooms and braise for another 5 minutes.

(c) *Articles*

There are three types of articles: definite article (*the*), indefinite article (*a* / *an*), and zero article. The treatment of articles is in fact one of the most difficult issues for a human translator working especially from Chinese into English. A machine translation system will find it difficult to change an indefinite article to definite article after a noun has been introduced or identified in the previous sentence.

I saw a girl. The girl was wearing a red dress.

Fortunately for translators, only the definite article is used most frequently in recipes. It is the addition of the definite article that worries the rule-writer of the machine translation system.

(d) *Conjunctions*

Conjunctions are words which are used to join clauses together, and to show the relationship between the ideas in the clauses. The functions of conjunctions in both English and Chinese are more or less the same. Nevertheless, conjunctions in English are used more frequently than in Chinese. This is because the clauses in a Chinese composite sentence are usually connected by parataxis, whereas those in an English complex or compound sentence are connected by hypotaxis. The main difference between parataxis and hypotaxis lies in the fact that connectives are much less imperatively needed in a Chinese composite sentence than in an English complex or compound sentence.

The paratactic and hypotactic differences between conjunctions in Chinese and those in English should also be considered together with the use of the set-off punctuation mark, known as *dun Hao* 頓號, in a Chinese text. A number of rules have been written to turn the Chinese paratactic sentence into a hypotactic structure.

將牛肉、豬肉、羊肉洗淨

wash clean the beef, pork, and mutton

將牛肉、豬肉洗淨切片

sauté beef and pork and shred

冬菇、蝦米浸透切粒

soak black mushrooms and shelled shrimps thoroughly and cut into dices.

(e) *Nouns*

Words that name persons, places, things, or ideas are nouns. Particular to recipes are nouns describing shapes. Recipes are full of nouns describing different shapes into which ingredients are cut.

界菱形花紋
slash surface with criss-cross pattern
切成方形
cut into squares
切成十字形花紋
cut into criss-cross patterns

The following example serves to illustrate the above point:

麵粉篩勻放入大碗中，慢慢加入滾水，迅速拌勻成柔軟粉糰，以少許麵粉爽手，將粉糰拍成長條形，再分切小圓粒，碾薄成圓形粉皮，放入肉丸做成小籠包形狀。

Sieve and place flour in a large bowl. Gradually add boiling water. Stir quickly to make a soft dough. Knead into rectangular roll. Dice and press into thin round sheets. Place on meatballs. Make into the shape of Shanghai pork dumplings.

(f) Prepositions

A preposition is a function word placed before nouns, pronouns, and gerunds to indicate a spatial or grammatical relationship. The following examples show that prepositions have to be handled contextually.

將蝦仁泡油 Soak shrimps in hot oil.
用滾水料滾水 Scald in boiling water with scalding ingredients.

(g) Pronouns

In Chinese, as in English, distinctions are made among he 他, she 她, it 它/牠, and he 牠. The agreement between nouns and pronouns is generally considered as one of the most difficult issues to handle since reference cannot always be found syntactically. Fortunately, recipes do not or need not contain pronouns which make the work of translation much easier.

雞洗淨，抹乾水斬件，加醃料醃十分鐘，泡油。
Wash the chicken clean, wipe dry, cut into pieces, marinate for 10 minutes and scald in oil.

(h) Verbs

In translating verbs, a number of issues, such as tense, aspect, and voice, have to be considered. Of these, tense is most important. As we all know, tense is the grammatical expression of time and is divided into the present and the past. There is no formal tense to express future action, which is determined by modality rather than the projected time of the action. It should be noted

that firstly, the concept of time and the grammatical tense belong to two different categories. Time exists independent of tense, which varies from language to language according to the linguistic expression of temporal relations. Secondly, tense forms in English are shown by variations in the morphological form of the verb, e.g. *I walk* and *I walked*, while the word forms of Chinese do not change, tense being expressed by adverbs of time 時間副詞, such as *xianzai* 現在(now), *jianglai* 將來 (in future), and *cengjing* 曾經 (once) and auxiliaries 助動詞 such as *liao* 了, *zhe* 着, and *guo* 過.

In translating recipes, only the present tense is used. Since they are written in directive format, agreement between noun and verb is avoided and all verbs are in their basic forms.

拖水後瀝乾水分待用 scald, drain and set aside for later use.

Recipes: Linguistic Analyses

Recipes are not very different from some directives, basic operating instructions, and manuals where the terminology is relatively limited and the grammatical structure comparatively simple. Linguistically, recipes belong to the domain of procedural discourse. The main function of procedural discourse is to prescribe or tell the reader of the text the process of doing something. A typical text of a Chinese recipe and its translation (Yam, 1997: 82) is given below to illustrate the linguistic characteristics of procedural discourse.

〈生爆鱔片〉

做法:

- (1) 鱔魚去骨洗淨，用乾布抹淨，切成塊。
- (2) 將鱔魚放入碗中，用二湯匙粟粉撈勻，放入熱油中炸透，撈起瀝乾油，盛上碟。
- (3) 用約一湯匙油，爆蒜肉，倒下調味料，煮滾後，淋上鱔魚面，再放上芫荽，即可供食。

Deep-fried eel slices

Method:

- (1) Remove bones of eel and wash thoroughly. Wipe dry and slice.
- (2) Put eel in casserole and mix with 2 tbsp cornflour. Fry in hot oil, drain and serve.
- (3) Saute garlic with 1 tbsp oil. Stir in seasonings and bring to boil. Pour sauce over eels and garnish with parsley. Serve hot.

It can be seen from the above that recipes are directives, which are a type of imperative sentence. What are the characteristics of a directive sentence? According to *A Comprehensive Grammar of the English Language* (Quirk et al., 1985: 827), there are two special features. Firstly, it generally has no

subject. Secondly, it usually has a main verb in the based form. Also, there are seven types of directives:

- V: Jump.
- VO: Open the door.
- VC: Be reasonable.
- VA: Get inside.
- VOO: Tell me the truth.
- VOC: Consider yourself lucky.
- VOA: Put the flowers on the table.

What is more, “the imperative verb lacks tense distinction and does not allow modal auxiliaries. The progressive form is rare, and the perfective even rarer.”

When taken together, all of these features make them pliable to machine translation. Firstly, directive clauses are short, and this makes it easier to write syntactical rules for text generation with high accuracy. Secondly, all sentences are in simple present tense, which means the difficulty of having to work out the right tense for Chinese sentences is removed. Thirdly, as no subject is involved, the obstacle of noun and pronoun agreement in Chinese-English translation no longer exists. For this type of text, machine translation can be of sufficient quality to replace human translation. The only other factor that needs to be considered is that Cantonese cuisine has a large number of Cantonese expressions which should be dealt with computationally.

Constructing the *TransRecipe* System

The first stage in the construction of this cookbook translation system is to produce its algorithm by simulating the working process of a human translator. The prevalent programming language Visual Basic was used to build up the algorithm which combines the corpus-based, example-based, pattern-based, and rule-based approaches. For the generation of a target text, the input source text goes through four stages of processing: (1) The first stage is “example-based” processing where a database of examples in the source text language will be aligned with input sentences through an intralingual matching engine, and then translations of the aligned sentences will be produced by an interlingual matching engine. (2) The second stage is “corpus-based” processing where bilingual matching of individual words and phrases takes place between the source sentences and the machine-tractable general and specialised dictionaries. (3) The third stage is “pattern-based” processing where sentences with similar patterns will be matched bilingually. (4) The last stage is “rule-based” processing where rules for text generation will produce the translation of the source text. The following is the flowchart for *TransRecipe*:

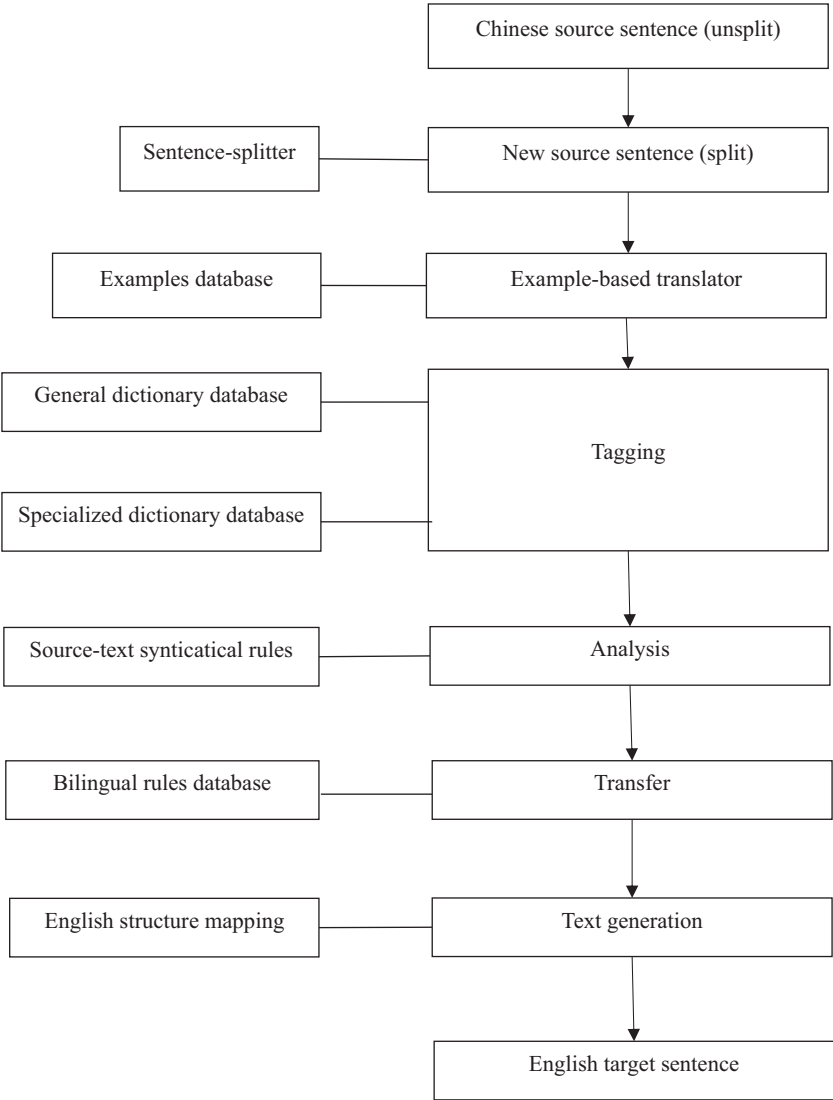


Figure 2.1 Flowchart of *TransRecipe*.

Stage 1: Example-based Machine Translation (EBMT)

This is the most linguistic of the corpus-based approaches in machine translation. Example translations are also used as the basis of new translations. The idea was first suggested by Makoto Nagao in 1984 in his article published

under the title “A Framework of a Mechanical Translation between Japanese and English by Analogy Principle”. (Nagao, 1984). He applied this method in his translation system and described it as “translating through analogy”. It was based on the idea that advances in computer technology have made it possible to gain access to the huge corpora of previously translated analogous examples to allow matching of bilingual expressions. It should be noted, however, that example-based machine translation cannot be wholly used for text generation as it is virtually impossible to translate every sentence by this approach. It is mainly a companion module to improve the overall quality of machine translation by serving as an alternative to the knowledge-based approach, and as an aid to the traditional rule-based methods. In other words, it should be used in conjunction with other empirical approaches to machine translation. A typical example of example-based machine translation is the Pangloss Mark III machine translation system.

With the use of examples, rules are not required. If the input sentence is aligned with a matching sentence, correct translation will be produced. If there is no matching, then correction is needed, and the new translation will serve as a new example in the bilingual corpus. This approach, needless to say, is dependent on the storage of a large number of translation examples in the database to enable the matching process. It is generally held that EBMT will include a sentence-aligned corpus, then there is the intralanguage matching between a stored source language chunk with an input chunk, the interlanguage matching between the source language chunk with the target language chunk, and finally there is the chunk combination.

In my construction of the Chinese-English *TransRecipe* system, the example-based approach has been used at the first stage. At this stage, a database of examples in the source text language has been aligned with input sentences through an intralingual matching engine, and then translations of the aligned sentences will be produced by an interlingual matching engine. Stored in this database is a bilingual corpus of example translations so that the source phrases or sentences will be translated by the best-match algorithm built up in the system. The example-based approach is best suited for frequently used expressions which cannot be easily dealt with by rules. For example, 芹菜絲、甘筍絲各少許 can be translated as “some shredded spring onion and a pick of flower-shaped carrot” and 麻油、胡椒粉各少許 can be rendered into “a dash of sesame oil and a pinch of pepper.” This approach saves the trouble of having to write a large number of rules to handle the collocational issues relating to the matching of different meanings of *shaoxu* 少許 with the different cooking ingredients used.

胡椒粉、麻油各少許 a pinch of pepper and sesame oil

胡椒粉、麻油少許 a pinch of pepper and sesame oil

胡椒粉、酒各少許 a pinch of pepper and some wine

胡椒粉、生粉各少許 a pinch of pepper and cornflour

胡椒粉、熟油各少許 a pinch of pepper and cooked oil

胡椒粉、糖各少許 a dash each of pepper and sugar
 薑片、陳皮各少許 some ginger slices and dried tangerine peel
 薑片、蔥段各少許 some ginger slices and spring onion sections
 薑片、蔥段少許 some ginger slices and spring onion sections
 薑片、紹酒、糖、生粉各少許 a little ginger juice, Shaoxing wine, sugar
 and cornflour

Stage 2: Corpus-based Machine Translation

This is the use of a pre-existing large corpus of lexical items or translated texts to construct a machine translation system. It is therefore closely related to the building of the lexical databases of the machine translation system. For *TransRecipe*, two lexical databases have been built: a “General Dictionary” and a “Specialised Dictionary.”

(a) *The General Dictionary*

Frequently used terms of different parts of speech are stored in the General Dictionary arranged according to the Hanyu Pinyin alphabetical order. It is expected that over 3,000 lexical items will have to be included in order to make the translation system work. Collected in this database are commonly-used expressions in Chinese cookbooks, such as 少量的 “a small amount of” (adjective), 均勻地 “evenly” (adverb), 及 “and” (conjunction), 5 分鐘 “5 minutes” (noun), 一切為二 “halve” (verb), and 罐 “can” (measure-word).

(b) *The Specialised Dictionary*

As recipes are normally made up of dish-names, ingredients and methods of preparation and cooking, the bilingual specialised dictionary database will be formed by terms in these three areas. The names of national and regional dishes form a major part of the bilingual database and terms related to the preparation of dishes have been collected and translated. About 10,000 items of Chinese food and drink and their translations, classified into 112 categories and given grammatical tags, have been stored in the database. Names and translations of dishes and soups, meat, bean curd, fish, shellfish, vegetables, and poultry, dim sum, ingredients, condiments, sauces, desserts, rice, noodles, congee, kitchen equipment, tableware, and restaurants, have been included in the database. Problems that have been encountered in the compilation of this database include the treatment of countable and uncountable nouns, the indication of measure words for all the common and proper nouns, the handling of variant characters with the same semantic meaning (e.g. *shuifen* as 水份 and 水分), the use of simplified and standard characters (*sa shang* as 洒上 and 灑上) in the source text, and the translation of different expressions bearing the



Figure 2.2 A food and drinks translation system.

same meaning (*zhu gun* 煮滾, *zhu zhi gun* 煮至滾 and *zhu zhi gun qi* 煮至滾起 “bring to the boil”).

To make this Specialised Dictionary readily accessible to the public, it has now been made into a CD-ROM in which a number of search engines have been designed to facilitate the search of a specific entry.

Stage 3: Pattern-based Machine Translation

The use of sentence patterns to deal with sentences which vary only in a small number of variables, such as the direct object or objects of the verb, has become increasingly popular. With the addition of the “partial matching method”, pattern-based translation has greatly enhanced the quality of machine output.

- 浸軟 soak until tender
- 泡軟 soak until tender
- 粉絲浸軟 soak mungbean vermicelli until tender.
- 花生用油炸成金黃色 deep-fry peanuts in oil until golden.
- 火腿用溫水沖洗過 rinse ham in warm water.
- 牛肉用清水浸十分鐘 soak beef in water for ten minutes.

Stage 4: Rule-based Machine Translation

Rule-based Machine Translation (RBMT) is a relatively traditional method which depends greatly on the difficult and time-consuming work of preparing and maintaining a large number of rules and a huge amount of lexical information in the form of dictionaries, both general and specialised.

In translating between Chinese and English, this is used to convert the syntactic structures of Chinese sentences into the structures of the equivalent English sentences by the use of some reordering rules, which rearrange the words or characters in the source text in the order of the target text. The rule-based approach is therefore serving the important role of bridging the structures of the two languages.

Two types of rules have been written for the parser of the system: those treating the measure words in the source text and those dealing with grammatical structures.

(a) Measure-word Rules

Rules governing the translation of measure words have been written to deal with the items mentioned in the “ingredients” and “seasoning” parts of the recipe. According to the views of the linguist, a measure word, or 量詞, is a type of classifier in Chinese that is placed between a numeral and a noun to indicate the unit of the object, a linguistic feature that is characteristic of the Sino-Tibetan language family. This type of measure word, which refers mainly to the ways things are counted, is known collectively as “nominal measure words” 名量詞 or “object measure words” 物量詞, such as 一本書 (a book), 三頭牛 (three cows). The other general type of measure words is “verbal measure words” 動量詞, which refers to the way actions are counted, such as 說一遍 (say it again), and 走一趟 (go there once).

In translating recipes, only the nominal measure words need to be treated. Nominal measure words can generally be divided into (1) “weights and measures” 度量量詞, such as 尺 (foot) and 斤 (catty); (2) “general unit measures” 一般單位量詞, referring to the way things are normally counted, such as 一個人 (a person), 一張紙 (a sheet of paper); and (3) “container measures” 容器量詞, referring to the way things have been contained or held, e.g. 一碗飯 and 一桶水. Another way to classify nominal measure words is to put them into two categories based on their grammatical structures: (1) simple measure words 單詞量詞, such as 張 and 尺; and (2) compound measure words 複詞量詞, such as 人次 (person-time) and 架次 (sortie).

(i) General Measure Words

A large number of measure words fall into this category. As each Chinese common noun carries with it a measure word, the number of measure words to be treated in the database is expected to be huge. In reality, this is not the case as most weight measures have been used for individual nouns, which greatly minimize the need to translate a huge number of measure words for each common noun.

(ii) Container Measure Words

There are two types of container measure words: single container measure words, such as 杯 (cup), and compound container measure words, such as 茶匙 (tsp). These measure words differ from others in their countability. All containers are countable common nouns. Some may be expressed in a form which does not need to indicate its singularity or plurality, such as tsp or tbsp. Others may need entries in the lexical database to indicate their singular and plural forms.

There are various forms of expression of measure words:

上湯六杯	6 cups of stock
上湯或水六杯	6 cups of stock or water

(iii) Weight Measure Words

Weight measure words are mostly used in the “ingredient” and “seasoning” sections of a recipe. Weights in recipes are usually given in terms of the Chinese traditional system of *qian* 錢, *liang* 兩, and *jin* 斤 and the British system of ounces and pounds. These systems need to be converted into the metric system.

Several methods have been used to translate Chinese measure words.

(1) METHOD OF AMPLIFICATION

This method is applied to fill grammatical or textual gaps when translating from one language into another. As measure words, which are strictly speaking different from classifiers, do not exist in English, Chinese-English translation involves the insertion or omission of proper measure words before nouns.

薑四厚片	4 thick slices ginger
薑一厚片	1 thick slice ginger
紅蘿蔔二條	2 carrots
紅蘿蔔一條	1 carrot

20 Technology for Translating, Visualizing, and Generating Recipes

(2) METHOD OF DIRECT EQUIVALENCE

The method of direct equivalence is used for “container measures” 容器量詞 as 杯 is translated as “cup” and 茶匙 as “teaspoon”.

蛋汁三湯匙
3 tbsp whisked egg
蛋汁一湯匙
1 tbsp whisked egg

(3) METHOD OF CONVERSION

The method of conversion is used when turning *jin* 斤 and *liang* 兩, units in the Chinese traditional system, into either the metric system in terms of grams (g) or British system in terms of ounces (oz).

牛肉一斤			
600 g beef			
Chinese	Metric		
一兩	1 tael (<i>liang</i>)	=	40 g
一斤	1 catty (<i>jin</i>)	=	600 g
British	Metric		
一安士	1 oz	=	30 g
一磅	1 lb	=	450 g

(4) METHOD OF COLLOCATIONAL TRANSLATION

More problematic is the translation of the general measure words which may or may not be translated, depending on the collocates, and which may be translated into a number of equivalents, depending on the nouns with which measure words are associated.

伊麵二個	2 cakes E-fu noodles
伊麵一個	1 cake E-fu noodles
蒜二粒	2 cloves garlic
蒜一粒	1 clove garlic

(5) METHOD OF MULTIPLE TRANSLATIONS

Measure words have a wide range of collocations. Rules have to be written to contextualize the translation.

一棵芫荽 a sprig of parsley
一棵西蘭花 a stalk of cauliflower

(b) *Grammatical Rules*

Grammatical rules, other than the rearrangement of word order in translating the source text into English, involve the use of the following translation methods.

(1) *Method of Addition*

The addition of the conjunction “and”

The set-off punctuation mark (、) is usually converted into a comma when translated into English. But used between two items, it has to be turned into a conjunction, which has to be added to the rules.

再將牛肉、豬肉加入 add beef and pork

The addition of the definite article

將牛肉洗淨 wash the beef

The addition of conjunction and article

將牛肉整條洗淨瀝乾 wash clean and dry the beef.

(2) *Method of Omission*

Omission is a translation method which leaves some words in the original untranslated to achieve grammatical accuracy and idiomaticity in the translation. Omission is used in *TransRecipe* for the following purposes:

- (i) To achieve syntactic completeness after breaking down the sentences according to preset rules.

當飯熟後 cook the rice
(當 and 後 have not been translated)

- (ii) To achieve brevity when the sentence is procedurally comprehensible.

再將牛肉放入鑊
Put the beef into the pan
(再 and 將 have been omitted)

然後放入蜜桃煮約十分鐘
Put in honey peach and boil for 10 minutes.
(然後 has not been translated)

(3) *Method of Fronting*

As with all imperative sentences, verbs in their base form are always placed at the beginning of the sentence. Grammatical rules for recipes have therefore been written to front the first verb in the Chinese source sentence in the

22 Technology for Translating, Visualizing, and Generating Recipes

English translation. For Chinese sentences beginning with a verb, no fronting is necessary, as in:

放入牛肉、豬肉及羊肉煮至滾起
Put in beef, pork and mutton and bring to the boil.

But when the first verb in the source sentence is not in the initial position, then fronting is necessary, and rules should be written to put the translation into a proper imperative sentence in English. The following are Chinese imperative sentences that begin with different parts of speech and their first verbs have been fronted in the translations:

將青豆及磨菰飛水 (*adv.*)
scald green beans and button mushrooms
與蝦仁同用少許生粉撈勻 (*conj.*)
stir well shelled shrimps with some cornflour
蝦連殼洗淨 (*noun*)
wash the shrimps with shells
用小火將材料焗至熟 (*prep.*)
bake ingredients over low heat until cooked
其餘原隻蒸熟 (*pron.*)
steam the rest until cooked.

Stage 5: Testing the System

When the lexical and syntactical databases are in place, the system can be put into use. Around 500 recipes have so far been tested using the system, and it is estimated that 1,000 recipes have to be tested before *TransRecipe* can be put on the market. This testing process is extremely important for the evaluation of the system in the areas of accuracy, clarity, speed, and reliability, and also for improvements to be made on the effectiveness of the parser and the enhancement and optimization of the various databases.

Demonstrating the *TransRecipe* System

Source text

〈燴牛仔骨〉

材料:

牛仔骨四至五條, 洋蔥一個, 番茄二個, 蒜茸、乾蔥蔥茸各一茶匙

醃料:

生抽 $\frac{3}{4}$ 湯匙, 酒半湯匙, 胡椒粉少許

調味料:

茄汁三湯匙, 生抽1%茶匙, 鹽 $\frac{1}{3}$ 茶匙, 糖二茶匙, 水1%杯

做法:

- (1) 牛仔骨斬成大段，放入醃料拌勻。
- (2) 洋蔥去外衣，切成大塊，略泡油待用。
- (3) 洗淨番茄，切大塊，去籽後再洗淨，瀝去水分，待用。
- (4) 燒熱二湯匙油，爆香乾蔥茸、蒜茸，加入番茄塊略炒，放入洋蔥、調味料及牛仔骨煮至滾起。
- (5) 將上項滾起的材料用中火爐煮至材料腦及汁濃，灑入胡椒粉少許，即可供食。

TransRecipe output

Stewed short ribs

Ingredients:

4-5 beef short ribs. 1 onion. 2 tomatoes. 1 tsp each of mashed garlic and minced shallots

Seasoning:

3/4 tbsp light soy. 1/2 tbsp wine. A dash of pepper

Seasonings:

3 tbsp tomato sauce. 1 1/2 tsp light soy. 1/3 tsp salt. 2 tsp sugar. 1 1/2 cups water

Method:

- (1) Chop beef short ribs into sections. Marinate with seasonings.
- (2) Peel onions. Cut into large pieces. Fry briefly in hot oil and set aside.
- (3) Wash clean tomatoes. Cut into large pieces. Seed and wash clean. Drain. Set aside.
- (4) Heat 2 tbsp of oil. Sauté minced shallot and mashed garlic. Add tomato and stir-fry for a while. Place, onions, seasonings and beef short ribs and bring to the boil.
- (5) Place boiled ingredients over medium heat and stew until cooked and sauce thickens. Sprinkle with a dash of pepper. Serve.

We must, of course, bear in mind that while each system has its own strengths, few of them produce satisfactory results when applied to real world texts. Logically, a hybrid system would fare better than a single-approach system. And it is quite likely that domain-specific systems will prevail in the near future as general application or all-purpose systems have yet to demonstrate their workability in a variety of contexts.

Conclusion

It goes without saying that as we are now in the age of artificial intelligence, and translation is inevitably widely used to overcome cross-lingual problems in data extraction and consumption. Web translation, as a result, has become increasingly popular. The main obstacle that cannot be easily solved is that no programming language, no matter how powerful it is, is capable of fully

and sufficiently processing multifaceted information available on the Internet written in the major world languages. This means that while it is important to achieve globalisation with the use of web translation, considerable allowance must be given to its inadequacies and imperfections due to the inability of an artificial language to match with a natural language. But for domain-specific texts, matching the languages is possible as its limited vocabularies and syntactical variations allow for easier machine processing. The relatively smooth running of *TransRecipe* demonstrates that with the application of translation methods commonly used in translation practice, a new “translational approach” is an alternative to the traditional linguistic approach. With this new approach, we are ready to march on into other domains. And as we do so, a general application system will eventually emerge. Then, and only then, can machine translation remove the linguistic and cultural barriers that have so far hindered global communication.

The following five chapters will provide detailed information on the points succinctly covered in this chapter.

References

- Batra, Devansh, Nirav Diwan, Utkarsh Upadhyay, Jushaan Singh Kalra, Tript Sharma, Aman Kumar Sharma, Dheeraj Khanna, Jaspreet Singh Marwah, Srilakshmi Kalathil, Navjot Singh *et al.* (2020) “Reciped: A Resource for Exploring Recipes”, *Database*, 2020.
- Bentivogli, Luisa, Arianna Bisazza, Mauro Cettolo, and Marcello Federico (2016) “Neural versus Phrase-based Machine Translation Quality: A Case Study”, *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 257–267.
- Bień, Michał, Michał Gilski, Martyna Maciejewska, Wojciech Taisner, Dawid Wisniewski, and Agnieszka Lawrynowicz (2020) “RecipeNLG: A Cooking Recipes Dataset for Semi-structured Text Generation”, *Proceedings of the 13th International Conference on Natural Language Generation*, 22–28.
- Borji, Ali (2022) “Generated Faces in the Wild: Quantitative Comparison of Stable Diffusion, Midjourney and Dall-e 2”, *arXiv preprint arXiv:2210.00586*.
- Bosch, Marc, Fengqing Zhu, Nitin Khanna, Carol J. Boushey, and Edward J. Delp (2011) “Combining Global and Local Features for Food Identification in Dietary Assessment”, *2011 18th IEEE International Conference on Image Processing*, 1789–1792.
- Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell *et al.* (2020) “Language Models are Few-shot Learners”, *Advances in Neural Information Processing Systems*, 33: 1877–1901.
- Catford, J.C. (1965) *A Linguistic Theory of Translation: An Essay in Applied Linguistics*, London: Oxford University Press.
- Chan, Sin-wai (2002) “The Making of *TransRecipe*: A Translational Approach to the Machine Translation of Chinese Cookbooks”, in Chan Sin-wai (ed.), *Translation and Information Technology*, Hong Kong: The Chinese University Press, 3–22.
- Chan, Sin-wai (ed.) (2002) *Translation and Information Technology*, Hong Kong: The Chinese University Press.
- Chandu, Khyathi, Eric Nyberg, and Alan W. Black (2019) “Storyboarding of Recipes: Grounded Contextual Generation”, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 6040–6046.
- Chang, Minsuk, Léonore V. Guillain, Hyeunghik Jung, Vivian M. Hare, Juho Kim, and Maneesh Agrawala (2018) “Recipeescape: An Interactive Tool for Analyzing

- Cooking Instructions at Scale”, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12.
- Chen, Jing-jing and Chong-Wah Ngo (2016) “Deep-based Ingredient Recognition for Cooking Recipe Retrieval”, *Proceedings of the 24th ACM International Conference on Multimedia*, 32–41.
- Chen, Jing-Jing, Chong-Wah Ngo, Fu-Li Feng, and Tat-Seng Chua (2018) “Deep understanding of cooking procedure for cross-modal recipe retrieval” *Proceedings of the 26th ACM International Conference on Multimedia*, 1020–1028.
- Chen, Jing-jing, Chong-Wah Ngo, and Tat-Seng Chua (2017) “Cross-modal Recipe Retrieval with Rich Food Attributes”, *Proceedings of the 25th ACM International Conference on Multimedia*, 1771–1779.
- Chen, Mei, Kapil Dhingra, Wen Wu, Lei Yang, Rahul Sukthankar, and Jie Yang (2009) “PFID: Pittsburgh Fast-food Image Dataset”, *2009 16th IEEE International Conference on Image Processing (ICIP)*, 289–292.
- Chen, Mei-Yun, Yung-Hsiang Yang, Chia-Ju Ho, Shih-Han Wang, Shane-Ming Liu, Eugene Chang, Che-Hua Yeh, and Ming Ouhyoung (2012) “Automatic Chinese Food Identification and Quantity Estimation”, *SIGGRAPH Asia 2012 Technical Briefs*, 1–4.
- Chen, Nicholas, Yun Young Lee, Maurice Rabb, and Bruce Schatz (2010) “Toward Dietary Assessment via Mobile Phone Video Cameras”, *AMIA Annual Symposium Proceedings*, American Medical Informatics Association, 2010, 106.
- Cho, Jaemin, Abhay Zala, and Mohit Bansal (2022) “DALL-Eval: Probing the Reasoning Skills and Social Biases of Text-to-image Generative Transformers”, *arXiv preprint arXiv:2202.04053*.
- Cho, Kyunghyun, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio (2014) “Learning Phrase Representations Using RNN Encoder–decoder for Statistical Machine Translation”, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734.
- Cho, Kyunghyun, Bart Van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio (2014) “On the Properties of Neural Machine Translation: Encoder-decoder Approaches”, *arXiv preprint arXiv:1409.1259*.
- Christiano, Paul F., Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei (2017) “Deep Reinforcement Learning from Human Preferences”, *Advances in Neural Information Processing Systems*, 30.
- Church, Kenneth W. and William A. Gale (1991) “Concordances for Parallel Texts”, *Using Corpora: Proceedings of the 7th Annual Conference of the UW for the New OED and Text Research*, Oxford: Oxford University Press, 40–62.
- Deeney, John J. (1995) “Transcription, Romanization, Transliteration”, in Chan Sin-wai and David E. Pollard (eds.), *An Encyclopaedia of Translation: Chinese-English. English-Chinese*, Hong Kong: The Chinese University Press, 1085–1097.
- DeFrancis, John (1996) *ABC Chinese-English Dictionary*, Hong Kong: The Chinese University Press.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018) “Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding”, *arXiv preprint arXiv:1810.04805*.
- Fatemi, Bahare, Quentin Duval, Rohit Girdhar, Michal Drozdal, and Adriana Romero-Soriano (2023) “Learning to Substitute Ingredients in Recipes”, *arXiv preprint arXiv:2302.07960*.

- Fei, Hongliang, Tan Yu, and Ping Li (2021) “Cross-lingual Cross-modal Pretraining for Multimodal Retrieval”, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 3644–3650.
- Frome, Andrea, Greg S. Corrado, Jonathon Shlens, Samy Bengio, Jeffrey Dean, Marc’Aurelio Ranzato, and Tomas Mikolov (2013) “DeViSE: A Deep Visualesemantic Embedding Model”, *Proceedings of the 26th International Conference on Neural Information Processing Systems-Volume 2*, 2121–2129.
- Fu, Han, Rui Wu, Chenghao Liu, and Jianling Sun (2020) “Mcen: Bridging cross-modal gap between cooking recipes and dish images with latent variable model” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14570–14580.
- Fujii, Tatsuki, Yuichi Sei, Yasuyuki Tahara, Ryohei Orihara, and Akihiko Ohsuga (2019) “‘Never Fry Carrots without Chopping’ Generating Cooking Recipes from Cooking Videos Using Deep Learning Considering Previous Process”, *International Journal of Networked and Distributed Computing*, 7(3): 107–112.
- Fung, Kam Ling 馮金陵, Lee Ngan Woon 李銀煥, and Hui Choi Yip 許彩葉 (1994) 《清爽涼拌》 (*Cold Dishes*), Hong Kong: Food Paradise Publishing Co. 飲食天地出版社.
- Gabeur, Valentin, Arsha Nagrani, Chen Sun, Karteek Alahari, and Cordelia Schmid (2022) “Masking Modalities for Cross-modal Video Retrieval”, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 1766–1775.
- Galatolo, Federico A., Mario G.C.A. Cimino, and Edoardo Cogotti (2022) “TeTImEval: A Novel Curated Evaluation Data Set for Comparing Text-to-image Models”, *arXiv preprint arXiv:2212.07839*.
- Gatt, Albert and Emiel Kraemer (2018) “Survey of the State of the Art in Natural Language Generation: Core Tasks, Applications and Evaluation”, *Journal of Artificial Intelligence Research*, 61: 65–170.
- Ghoddosian, Reza, Saif Sayed, and Vassilis Athitsos (2022) “Hierarchical Modeling for Task Recognition and Action Segmentation in Weakly-labeled Instructional Videos”, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 1922–1932.
- Goel, Mansi, Pallab Chakraborty, Vijay Ponnaganti, Minnet Khan, Sritanaya Tatipamala, Aakanksha Saini, and Ganesh Bagler (2022) “Ratatouille: A Tool for Novel Recipe Generation”, *2022 IEEE 38th International Conference on Data Engineering Workshops (ICDEW)*, 107–110.
- Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David WardeFarley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio (2014) “Generative Adversarial Networks”, *arXiv preprint arXiv:1406.2661*.
- Gove, Philip Babcock (comp.) (1986) *Webster’s Third New International Dictionary*, Springfield, Mass.: Merriam-Webster Inc.
- Guerrero, Ricardo, Hai X. Pham, and Vladimir Pavlovic (2021) “Cross-modal Retrieval and Synthesis (X-MRS): Closing the Modality Gap in Shared Subspace Learning”, *Proceedings of the 29th ACM International Conference on Multimedia*, 3192–3201.
- Guo, Zhao, Lianli Gao, Jingkuan Song, Xing Xu, Jie Shao, and Heng Tao Shen (2016) “Attention-based LSTM with Semantic Consistency for Videos Captioning”, *Proceedings of the 24th ACM International Conference on Multimedia*, 357–361.

- Han, Fangda, Guoyao Hao, Ricardo Guerrero, and Vladimir Pavlovic (2021) “Multi-attribute Pizza Generator: Cross-domain Attribute Control with Conditional StyleGAN”, *arXiv preprint arXiv:2110.11830*.
- Harashima, Jun, Yuichiro Someya, and Yohei Kikuta (2017) “Cookpad Image Dataset: An Image Collection as Infrastructure for Food Research”, *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1229–1232.
- Hasyim, Muhammad, Firman Saleh, Rudy Yusuf, and Asriani Abbas (2021) “Artificial Intelligence: Machine Translation Accuracy in Translating FrenchIndonesian Culinary Texts”, *International Journal of Advanced Computer Science and Applications*, 12 (3).
- Haynes, Colin (1998) *Breaking down the Language Barriers*, London: Aslib.
- Hearst, Marti A., Susan T. Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf (1998) “Support Vector Machines”, *IEEE Intelligent Systems and Their Applications*, 13(4): 18–28.
- Hertz, Amir, Ron Mokady, Jay Tenenbaum, Kfir Aberman, Yael Pritch, and Daniel Cohen-Or (2022) “Prompt-to-prompt Image Editing with Cross Attention Control”, *arXiv preprint arXiv:2208.01626*.
- Heusel, Martin, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter (2017) “GANs Trained by a Two Time-scale Update Rule Converge to a Local Nash Equilibrium”, *Advances in Neural Information Processing Systems*, 30.
- Ho, Jonathan, Ajay Jain, and Pieter Abbeel (2020) “Denosing Diffusion Probabilistic Models”, *Advances in Neural Information Processing Systems*, 33: 6840–6851.
- Hochreiter, Sepp (1998) “The Vanishing Gradient Problem during Learning Recurrent Neural Nets and Problem Solutions”, *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, 6(2): 107–116.
- Hochreiter, Sepp and Jürgen Schmidhuber (1997) “Long Short-term Memory”, *Neural Computation*, 9(8): 1735–1780.
- Horita, Daichi, Ryosuke Tanno, Wataru Shimoda, and Keiji Yanai (2018) “Food Category Transfer with Conditional CycleGAN and a Large-scale Food Image Dataset”, *Proceedings of the Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management*, 67–70.
- Hu, Fu 胡附 (1984) 《數詞和量詞》 (*Numerals and Measure-words*), Shanghai: Shanghai Education Press 上海教育出版社.
- Hu, Yuheng, Lydia Manikonda, and Subbarao Kambhampati (2014) “What We Instagram: A First Analysis of Instagram Photo Content and User Types”, *Proceedings of the International AAAI Conference on Web and Social Media*, 8: 595–598.
- Huang, Po-Yao, Xiaojun Chang, Alexander Hauptmann, and Eduard Hovy (2020) “Forward and Backward Multimodal NMT for Improved Monolingual and Multilingual Cross-modal Retrieval”, *Proceedings of the 2020 International Conference on Multimedia Retrieval*, 53–62.
- Ifrah, Georges (2000) *The Universal History of Numbers: From Prehistory to the Invention of the Computer*, tr. David Bellos, E.F. Harding, Sophie Wood, and Ian Monk, New York: Wiley.
- Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros (2017) “Image-to-image Translation with Conditional Adversarial Networks”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1125–1134.
- Ji, Lei, Chenfei Wu, Daisy Zhou, Kun Yan, Edward Cui, Xilin Chen, and Nan Duan (2022) “Learning Temporal Video Procedure Segmentation from an Automatically

- Collected Large Dataset”, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 1506–1515.
- Kagaya, Hokuto, Kiyoharu Aizawa, and Makoto Ogawa (2014) “Food Detection and Recognition Using Convolutional Neural Network”, *Proceedings of the 22nd ACM International Conference on Multimedia*, 1085–1088.
- Kawano, Yoshiyuki and Keiji Yanai (2014) “Foodcam: A Real-time Mobile Food Recognition System Employing Fisher Vector”, *MultiMedia Modeling: 20th Anniversary International Conference, MMM 2014, Dublin, Ireland, January 6-10, 2014, Proceedings*, Part II 20, 369–373.
- Kawar, Bahjat, Shiran Zada, Oran Lang, Omer Tov, Huiwen Chang, Tali Dekel, Inbar Mosseri, and Michal Irani (2022) “Imagic: Text-based Real Image Editing with Diffusion Models”, *arXiv preprint arXiv:2210.09276*.
- Khan, Rijwan, Santosh Kumar, Niharika Dhingra, and Neha Bhati (2021) “The Use of Different Image Recognition Techniques in Food Safety: A Study”, *Journal of Food Quality*, 1–10.
- Kiddon, Chloé, Luke Zettlemoyer, and Yejin Choi (2016) “Globally Coherent Text Generation with Neural Checklist Models”, *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 329–339.
- Kitamura, Keigo, Toshihiko Yamasaki, and Kiyoharu Aizawa (2008) “Food Log by Analyzing Food Images”, *Proceedings of the 16th ACM International Conference on Multimedia*, 999–1000.
- Kitamura, Keigo, Toshihiko Yamasaki, and Kiyoharu Aizawa (2009) “Foodlog: Capture, Analysis and Retrieval of Personal Food Images via Web”, *Proceedings of the ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities*, 23–30.
- Koehn, Philipp and Rebecca Knowles (2017) “Six Challenges for Neural Machine Translation”, *Proceedings of the First Workshop on Neural Machine Translation*, 28–39.
- Koehn, Philipp, Franz Josef Och, and Daniel Marcu (2003) “Statistical Phrase-based Translation”, *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 127–133.
- Lee, Helena, Ke Shu, Palakorn Achananuparp, Philips Kokoh Prasetyo, Yue Liu, Ee-Peng Lim, and Lav R. Varshney (2020) “RecipeGPT: Generative Pretraining Based Cooking Recipe Generation and Evaluation System”, *Companion Proceedings of the Web Conference 2020*, 181–184.
- Lee, Ngan Woon 李銀煥 (1996) 《家常老幼靚湯》 (*Soup for the Whole Family*), Hong Kong: Sea Shore Book Company 海濱圖書公司.
- Li, Guang, Shubo Ma, and Yahong Han (2015) “Summarization-based Video Caption via Deep Neural Networks”, *Proceedings of the 23rd ACM International Conference on Multimedia*, 1191–1194.
- Li, Jiatong, Fangda Han, Ricardo Guerrero, and Vladimir Pavlovic (2021) “Picture-to-amount (pita): Predicting Relative Ingredient Amounts from Food Images”, *2020 25th International Conference on Pattern Recognition (ICPR)*, 10343–10350.
- Liu, Chang, Yu Cao, Yan Luo, Guanling Chen, Vinod Vokkarane, and Yunsheng Ma (2016) “Deepfood: Deep Learning-based Food Image Recognition for Computer-aided Dietary Assessment”, *Inclusive Smart Cities and Digital Health: 14th International Conference on Smart Homes and Health Telematics, ICOST 2016, Wuhan, China, May 25-27, 2016. Proceedings 14*, 37–48.
- Liu, Jiatong (2021) “Multimodal Machine Translation”, *IEEE Access*.
- Liu, Vivian and Lydia B. Chilton (2022) “Design Guidelines for Prompt Engineering Text-to-image Generative Models”, *CHI Conference on Human Factors in Computing Systems*, 1–23.

- Liu, Xiao, Yansong Feng, Jizhi Tang, Chengang Hu, and Dongyan Zhao (2022) “Counterfactual Recipe Generation: Exploring Compositional Generalization in a Realistic Scenario”, *arXiv preprint arXiv:2210.11431*.
- Liu, Zhiming, Kai Niu, and Zhiqiang He (2023) “ML-CookGAN: Multi-label Generative Adversarial Network for Food Image Generation”, *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(2s): 1–21.
- Majumder, Bodhisattwa Prasad, Shuyang Li, Jianmo Ni, and Julian McAuley (2019) “Generating Personalized Recipes from Historical User Preferences”, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 5976–5982.
- Malmaud, Jonathan, Jonathan Huang, Vivek Rathod, Nicholas Johnston, Andrew Rabinovich, and Kevin Murphy (2015) “‘What’s Cooking?’ Interpreting Cooking Videos using Text, Speech, and Vision” *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 143–152.
- Marin, Javier, Aritro Biswas, Ferda Ofli, Nicholas Hynes, Amaia Salvador, Yusuf Aytar, Ingmar Weber, and Antonio Torralba (2019) “Recipe1m+: A Dataset for Learning Cross-modal Embeddings for Cooking Recipes and Food Images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(1): 187–203.
- Metz, Cade and Priya Krishna (2022) “Can AI Write Recipes Better Than Humans?” *The Seattle Times* www.seattletimes.com/business/can-ai-write-recipes-better-than-humans/, retrieved on 17 April 2023.
- Min, Weiqing, Shuqiang Jiang, Jitao Sang, Huayang Wang, Xinda Liu, and Luis Herranz (2016) “Being a Supercook: Joint Food Attributes and Multimodal Content Modeling for Recipe Retrieval and Exploration”, *IEEE Transactions on Multimedia*, 19(5): 1100–1113.
- Mirza, Mehdi and Simon Osindero (2014) “Conditional Generative Adversarial Nets”, *arXiv preprint arXiv:1411.1784*.
- Mohammadshahi, Alireza, Rémi Lebret, and Karl Aberer (2019) “Aligning Multilingual Word Embeddings for Cross-modal Retrieval Task”, *Proceedings of the Beyond Vision and LANGUAGE: inTEgrating Real-world kNOWLEDge (LANTERN)*, 11–17.
- Mokady, Ron, Amir Hertz, Kfir Aberman, Yael Pritch, and Daniel Cohen-Or (2022) “Null-text Inversion for Editing Real Images Using Guided Diffusion Models”, *arXiv preprint arXiv:2211.09794*.
- Mori, Shinsuke, Hirokuni Maeta, Tetsuro Sasada, Koichiro Yoshino, Atsushi Hashimoto, Takuya Funatomi, and Yoko Yamakata (2014) “Flowgraph2text: Automatic Sentence Skeleton Compilation for Procedural Text Generation”, *Proceedings of the 8th International Natural Language Generation Conference (INLG)*, 118–122.
- Mori, Shinsuke, Hirokuni Maeta, Yoko Yamakata, and Tetsuro Sasada (2014) “Flow Graph Corpus from Recipe Texts”, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC’14)*, 2370–2377.
- Nagao, Makoto (1984) “A Framework of a Mechanical Translation between Japanese and English by Analogy”, in Alick Elithorn and Ranan Barneiji (eds.), *Artificial and Human Intelligence*, Amsterdam: North-Holland Publishing Company, 73–80.
- Newmark, Peter (1981) *Approaches to Translation*, Oxford: Pergamon Press.
- Newmark, Peter (1988) *A Textbook of Translation*, Hertfordshire: Prentice-Hall.
- Nichol, Alexander Quinn, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew, Ilya Sutskever, and Mark Chen (2022) “GLIDE: Towards

- Photorealistic Image Generation and Editing with TextGuided Diffusion Models”, *International Conference on Machine Learning*, 16784–16804.
- Nishimura, Taichi, Atsushi Hashimoto, and Shinsuke Mori (2019) “Procedural Text Generation from a Photo Sequence”, *Proceedings of the 12th International Conference on Natural Language Generation*, 409–414.
- OpenAI (2022) “Introducing ChatGPT”, accessed on 26 April 2023, available at <https://openai.com/blog/chatgpt>.
- Pan, Liang-Ming, Jingjing Chen, Jianlong Wu, Shaoteng Liu, Chong-Wah Ngo, Min-Yen Kan, Yugang Jiang, and Tat-Seng Chua (2020) “Multi-modal Cooking Workflow Construction for Food Recipes”, *Proceedings of the 28th ACM International Conference on Multimedia*, 1132–1141.
- Papadopoulos, Dim P., Youssef Tamaazousti, Ferda Ofli, Ingmar Weber, and Antonio Torralba (2019) “How to Make a Pizza: Learning a Compositional Layer-based Gan Model”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8002–8011.
- Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu (2002) “Bleu: A Method for Automatic Evaluation of Machine Translation”, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 311–318.
- Pavlichenko, Nikita, Fedor Zhdanov, and Dmitry Ustalov (2022) “Best Prompts for Text-to-image Models and How to Find Them”, *arXiv preprint arXiv:2209.11711*.
- Petsiuk, Vitali, Alexander E. Siemenn, Saisamrit Surbehera, Zad Chin, Keith Tyser, Gregory Hunter, Arvind Raghavan, Yann Hicke, Bryan A. Plummer, Ori Kerret *et al.* (2022) “Human Evaluation of Text-to-image Models on a Multi-task Benchmark”, *arXiv preprint arXiv:2211.12112*.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*, New York: Longman Group Ltd.
- Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever *et al.* (2019) “Language Models Are Unsupervised Multitask Learners”, *OpenAI Blog*, 1(8): 9.
- Radford, Alec, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark *et al.* (2021) “Learning Transferable Visual Models from Natural Language Supervision”, *International Conference on Machine Learning*, 8748–8763.
- Ramesh, Aditya, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen (2022) “Hierarchical Text-conditional Image Generation with Clip Latents”, *arXiv preprint arXiv:2204.06125*.
- Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer (2022) “High-resolution Image Synthesis with Latent Diffusion Models”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10684–10695.
- Sagart, Laurent (1999) *The Roots of Old Chinese*, Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Sahoo, Doyen, Wang Hao, Shu Ke, Wu Xiongwei, Hung Le, Palakorn Achananuparp, Ee-Peng Lim, and Steven C.H. Hoi (2019) “FoodAI: Food Image Recognition via Deep Learning for Smart Food Logging”, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2260–2268.
- Salvador, Amaia, Michal Drozdal, Xavier Giro-i Nieto, and Adriana Romero (2019) “Inverse Cooking: Recipe Generation from Food Images”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10453–10462.

- Salvador, Amaia, Erhan Gundogdu, Loris Bazzani, and Michael Donoser (2021) “Revamping Cross-modal Recipe Retrieval with Hierarchical Transformers and Self-supervised Learning”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15475–15484.
- Salvador, Amaia, Nicholas Hynes, Yusuf Aytar, Javier Marin, Ferda Ofli, Ingmar Weber, and Antonio Torralba (2017) “Learning Cross-modal Embeddings for Cooking Recipes and Food Images”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3020–3028.
- Sato, Takayuki, Jun Harashima, and Mamoru Komachi (2016) “Japanese-English Machine Translation of Recipe Texts”, *Proceedings of the 3rd Workshop on Asian Translation (WAT2016)*, 58–67.
- Schäffner, Christina and Helen Kelly-Holmes (1995) *Cultural Functions of Translation*, Clevedon: Multilingual Matters.
- Sennrich, Rico, Barry Haddow, and Alexandra Birch (2016) “Improving Neural Machine Translation Models with Monolingual Data”, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 86–96.
- Seo, Paul Hongsuck, Arsha Nagrani, Anurag Arnab, and Cordelia Schmid (2022) “End-to-end Generative Pretraining for Multimodal Video Captioning”, *arXiv preprint arXiv:2201.08264*.
- Simoons, Frederick J. (1991) *Food in China: A Cultural and Historical Inquiry*, Boca Raton: CRC Press.
- Sohl-Dickstein, Jascha, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli (2015) “Deep Unsupervised Learning Using Nonequilibrium Thermodynamics”, *International Conference on Machine Learning*, 2256–2265.
- Song, Yang, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole (2020) “Score-based Generative Modeling through Stochastic Differential Equations”, *arXiv preprint arXiv:2011.13456*.
- Su, Han, Ting-Wei Lin, Cheng-Te Li, Man-Kwan Shan, and Janet Chang (2014) “Automatic Recipe Cuisine Classification by Ingredients”, *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 565–570.
- Sugiyama, Yu and Keiji Yanai (2021) “Cross-modal Recipe Embeddings by Disentangling Recipe Contents and Dish Styles”, *Proceedings of the 29th ACM International Conference on Multimedia*, 2501–2509.
- Sulubacak, Umut, Ozan Caglayan, Stig-Arne Grönroos, Aku Rouhe, Desmond Elliott, Lucia Specia, and Jörg Tiedemann (2020) “Multimodal Machine Translation through Visuals and Speech”, *Machine Translation*, 34: 97–147.
- Sun, Chen, Austin Myers, Carl Vondrick, Kevin Murphy, and Cordelia Schmid (2019) “VideoBERT: A Joint Model for Video and Language Representation Learning”, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 7464–7473.
- Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le (2014) “Sequence to Sequence Learning with Neural Networks”, *Advances in Neural Information Processing Systems*, 27: 3104–3112.
- Tai, Kuo-Chung (1979) “The Tree-to-tree Correction Problem”, *Journal of the ACM (JACM)*, 26(3): 422–433.
- Tam, Ng Wai Fung 譚吳威鳳 (1997) 《超值營養食譜》 (*Healthy Money-saving Meals*), Hong Kong: Wan Li Book Co. Ltd 萬里機構 and Food Paradise Publishing Co. 飲食天地出版社.

- Tanno, Ryosuke, Daichi Horita, Wataru Shimoda, and Keiji Yanai (2018) “Magical Rice Bowl: A Real-time Food Category Changer”, *Proceedings of the 26th ACM International Conference on Multimedia*, 1244–1246.
- Teng, Chun-Yuen, Yu-Ru Lin, and Lada A. Adamic (2012) “Recipe Recommendation Using Ingredient Networks”, *Proceedings of the 4th Annual ACM Web Science Conference*, 298–307.
- Treffry, Diana (ed.) (1998) *Collins English Dictionary*, 4th edition. Glasgow: HarperCollins Publishers.
- Ueda, Mayumi, Mari Takahata, and Shinsuke Nakajima (2011) “Recipe Recommendation Method Based on User’s Food Preferences”, *Proceedings of the IADIS International Conference on E-Society*, 591–594.
- Van Asbroeck, Stephanie and Christophe Matthys (2020) “Use of Different Food Image Recognition Platforms in Dietary Assessment: Comparison Study”, *JMIR Formative Research*, 4(12): e15602.
- Varshney, Lav R., Florian Pinel, Kush R. Varshney, Debarun Bhattacharjya, Angela Schörgendorfer, and Y-M Chee (2019) “A Big Data Approach to Computational Creativity: The Curious Case of Chef Watson”, *IBM Journal of Research and Development*, 63(1): 7–11.
- Vasconcellos, Muriel (1996) *Recent Trends in Machine Translation*, London: Aslib.
- Venuti, Lawrence (1995) *The Translator’s Invisibility: A History of Translation*, London and New York: Routledge.
- Vinay, Jean-Paul and Jean Darbelnet (1954, 1995) *Comparative Stylistics of French and English: A Methodology for Translation*, Juan C. Sager and M.-J. Hamel (tr. and ed.), Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Wang, Hao, Doyen Sahoo, Chenghao Liu, Ee-peng Lim, and Steven C.H. Hoi (2019) “Learning Cross-modal Embeddings with Adversarial Networks for Cooking Recipes and Food Images”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11572–11581.
- Wang, Liping, Qing Li, Na Li, Guozhu Dong, and Yu Yang (2008) “Substructure Similarity Measurement in Chinese Recipes”, *Proceedings of the 17th International Conference on World Wide Web*, 979–988.
- Wang, Su, Honghao Gao, Yonghua Zhu, Weilin Zhang, and Yihai Chen (2019) “A Food Dish Image Generation Framework Based on Progressive Growing GANs”, *Collaborative Computing: Networking, Applications and Worksharing: 15th EAI International Conference, CollaborateCom 2019, London, UK, August 19-22, 2019, Proceedings 15*, 323–333.
- Wang, Wenjie, Ling-Yu Duan, Hao Jiang, Peiguang Jing, Xuemeng Song, and Liqiang Nie (2021) “Market2Dish: Health-aware Food Recommendation”, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 17(1): 1–19.
- Weellisch, Hans (1975) *Transcription and Transliteration*, Silver Spring, MD: Institute of Modern Language.
- Wiktionary (CC BY-SA 3.0) <https://en.wiktionary.org/wiki/recipe> accessed: 2023-04-27
- Xie, Haoran, Lijuan Yu, and Qing Li (2010) “A Hybrid Semantic Item Model for Recipe Search by Example”, *2010 IEEE International Symposium on Multimedia*, 254–259.
- Xie, Zhongwei, Ling Liu, Lin Li, and Luo Zhong (2021) “Learning Joint Embedding with Modality Alignments for Cross-modal Retrieval of Recipes and Food Images”, *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*, 2221–2230.

- Xu, Frank F., Lei Ji, Botian Shi, Junyi Du, Graham Neubig, Yonatan Bisk, and Nan Duan (2020) “A Benchmark for Structured Procedural Knowledge Extraction from Cooking Videos”, *Proceedings of the First International Workshop on Natural Language Processing Beyond Text*, 30–40.
- Yam, Lisa 方任利莎 (1997) 《方太食譜之魚蝦蟹》 (*Lisa Yam's Cook Book: Seafood*), Hong Kong: Ming Pao Press Ltd. 明窗出版社.
- Yamakata, Yoko, Shinji Imahori, Hirokuni Maeta, and Shinsuke Mori (2016) “A Method for Extracting Major Workflow Composed of Ingredients, Tools, and Actions from Cooking Procedural Text”, *2016 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 1–6.
- Yamamoto, Kohei and Keiji Yanai (2022) “Text-based Image Editing for Food Images with CLIP”, *Proceedings of the 7th International Workshop on Multimedia Assisted Dietary Management*, 29–37.
- Yanai, Keiji, Kaimu Okamoto, Tetsuya Nagano, and Daichi Horita (2019) “Large-scale Twitter Food Photo Mining and its Applications”, *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*, 77–85.
- Yang, Shulin, Mei Chen, Dean Pomerleau, and Rahul Sukthankar (2010) “Food Recognition Using Statistics of Pairwise Local Features”, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2249–2256.
- Yu, Zhiwei, Hongyu Zang, and Xiaojun Wan (2020) “Routing Enforced Generative Model for Recipe Generation”, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 3797–3806.
- Zhang, Deqin 張德鑫 (1999) 《數裏乾坤》 (*The Mysterious World of Numerals*), Beijing: Peking University Press 北京大學出版社.
- Zhang, Yixin, Yoko Yamakata, and Keishi Tajima (2019) “Categorization of Cooking Actions Based on Textual/Visual Similarity”, *Proceedings of the 5th International Workshop on Multimedia Assisted Dietary Management*, 42–49.
- Zhou, Kaiyang, Jingkang Yang, Chen Change Loy, and Ziwei Liu (2022) “Learning to Prompt for Vision-language Models”, *International Journal of Computer Vision*, 130(9): 2337–2348.
- Zhou, Luwei, Chenliang Xu, and Jason J. Corso (2018) “Towards Automatic Learning of Procedures from Web Instructional Videos”, *Thirty-second AAAI Conference on Artificial Intelligence*.
- Zhu, Bin, Chong-Wah Ngo, and Wing-Kwong Chan (2021) “Learning from Web Recipe-image Pairs for Food Recognition: Problem, Baselines and Performance”, *IEEE Transactions on Multimedia*, 24: 1175–1185.
- Zhu, Bin and Chong-Wah Ngo (2020) “CookGAN: Causality based text-to-image synthesis”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5519–5527.
- Zhu, Bin, Chong-Wah Ngo, Jingjing Chen, and Yanbin Hao (2019) “R2gan: Cross-modal Recipe Retrieval with Generative Adversarial Network”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11477–11486.
- Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A Efros (2017) “Unpaired Image-to-image Translation Using Cycle-consistent Adversarial Networks”, *Proceedings of the IEEE International Conference on Computer Vision*, 2223–2232.